# Perception in Real and Artificial Insects: A Robotic Investigation of Cricket Phonotaxis

Barbara Helen Webb

Ph.D.

University of Edinburgh
1993

## Abstract

The aim of this thesis is to investigate a methodology for studying perceptual systems by building artificial ones. It is proposed that useful results can be obtained from detailed robotic modelling of specific sensorimotor mechanisms in lower animals. By looking at the sensory control of behaviour in simple biological organisms, and in working robots, it is argued that proper appreciation of the physical interaction of the system with the environment and the task is essential for discovering how perceptual mechanisms function. Although links to biology, and concern with perceptual competence, are fields of growing interest in Artificial Intelligence, much of the current research fails to adequately address these issues, as the model systems being built do not represent real sensorimotor problems.

By analyzing what is required for a model of a system to contribute to explaining that system, a particular approach to modeling perceptual systems is suggested. This involves choosing an appropriate target system to model, building a system that validly represents the target with respect to a particular hypothesis, and properly evaluating the behaviour of the model system to draw conclusions about the target. The viability and potential contribution of this approach is demonstrated in the design, implementation and evaluation of a mobile robot model of a hypothesised mechanism for phonotaxis in the cricket.

The result is a robot that successfully locates a specific sound source under a variety of conditions, with a range of behaviour that resembles the cricket in many ways. This provides some support for the hypothesis that the neural mechanism for phonotaxis in crickets does not involve separate processing for recognition and location of the signal, as is generally supposed. It also shows the importance of understanding the physical interaction of the system's structure with its environment in devising and implementing perceptual systems. Both these results vindicate the proposed methodology.

## Acknowledgements

Many people have helped me to complete this work. It could not have been done without the support of my supervisors: Tim Smithers, who from the outset has encouraged and guided my attempts to explore new ideas and approaches; and John Hallam, whose expertise I have drawn upon constantly. The practical work described owes much to others, including Jim Donnett for the original Lego vehicle design, Sandy Colquhuon for implementing and maintaining the electronics, and Tim Colles for his work on the PSI processor, as well as the other technical staff in the A.I. department.

I would like to thank Dominic Hunter for his careful reading of the draft of this thesis, and also my other friends in Edinburgh for preventing me from losing my perspective. But most of all I would thank my husband, Chris, for his boundless patience in providing moral and physical support, from the decision to pursue a PhD to the final proof-reading of this text.

This work was supported by a British Commonwealth scholarship.

## Declaration

This thesis has been composed by myself and describes my own work.

From the fact that there are 400,000 species of beetles on this planet, but only 5000 species of mammals, he concluded that the Creator, if He exists, has a special preference for beetles ...

(Report of a lecture by J.B.S. Haldane, 1951)

# Table of Contents

# Chapter 1

# Introduction

## 1.1 Defining a question

This thesis is concerned with perceptual systems, that is, with the systems that underlie the sensory control of behaviour. This concept of perception differs from the traditional one—perception as the transformation of sensory input to conscious experience— but is adopted here for several reasons. It allows 'perceptual systems' in a wide range of biological and robotic systems to be discussed without making assumptions about the 'experiences' of these agents. It stresses the notion that the function of perception in controlling action is an essential determinant of the mechanisms of a perceptual system. It also removes the emphasis on 'internal representations' as the products or processes of perception.

The latter point places this approach to perception outside the mainstream of Artificial Intelligence research in which a major assumption about how intelligent systems work is that they use internal represention (Newell and Simon, 1976; Smith, 1985). In Chapter 2 I will argue that this assumption is undermined when the details of real perceptual systems are examined: in general, they don't work by processing representations (Varela *et al*, 1991; Brooks, 1991d). Looking at these systems also reinforces the importance of understanding perception in terms of behavioural control (Gibson, 1979).

To appreciate these ideas requires the investigation of *complete* perceptual systems, that is, examples where the whole path from sensors to behaviour can be examined. Understanding complete mechanisms can, currently, only be done

if the system is a relatively simple one, such as a lower animal (insect or below) or a mobile robot. AI has usually favoured the study of parts of complex systems rather than the whole of simple ones. This alternative approach can provide valuable insights into the nature of perceptual systems, that may help guide the investigation of more complex examples.

The capacity of an agent to interact competently with the surrounding environment, utilising sensing to successfully adapt actions to its situation, is a fundamental aspect of intelligent behaviour (Winograd and Flores, 1986). The idea that some research effort in AI should be directed towards understanding mechanisms of sensorimotor control in simple systems (and thus should interact with the study of biology rather than the study of cognition) has gained much support over the last five or so years (Langton, 1989; Meyer and Wilson, 1991). However, it still remains uncertain what methodology for investigating such systems is appropriate, and how the interaction with biology is best approached—"What is needed is a non-naive way of including biological insights. We believe that there are currently no generally accepted strategies available how to best structure the interaction between robot design and biology" (Pfeifer and Verschure, 1992). Thus the question this thesis attempts to answer is—what is an effective methodology for linking biological investigation to constructing artificial agents so as to obtain better understanding of perceptual systems?

## 1.2 Developing an answer

To answer this question requires some analysis of the purpose of building artificial systems as a means of investigating hypotheses, and the potential role for AI in investigating biological perceptual systems in particular. In Chapter 3, I provide such an analysis, casting AI as a process of modeling, and identifying important aspects of the function of model-building in science and engineering. In particular, I point out that the value of a model lies in the adequacy with which it represents the problem—in other words, the way in which a model represents its object must

be well specified for conclusions about its object to be drawn from operations on the model.

This analysis provides a basis for the evaluation of current approaches to simple perceptual systems in AI. Objections are raised against over-generalisation of the 'problem' of perception, over-simplification of the dynamic interaction of agents with the environment, and over-use of unstructured learning and evolution— on the basis that real perceptual systems are not adequately represented in such approaches.

By contrast, I propose that perceptual systems could be usefully investigated in AI by building and testing relatively detailed robotic models of particular biological sensorimotor mechanisms. This methodological proposal is the basis of the experimental work described in this thesis, which is aimed at establishing the viability and potential contribution of this approach.

The proposal has three main components: choosing a particular biological system as a target and modeling it in sufficient detail to reflect hypotheses about the mechanism; using a physical robotic model to capture equivalent problems of interacting with the real world; and carrying out experimental tests to examine properly the hypotheses embedded in the model.

### 1.2.1 A target system

If simple animals are appropriate target systems, neuroethology can provide a rich field of examples. Its aims—to understand neural mechanisms underlying behaviour (Camhi, 1987)—are closely related to the problem of understanding perception discussed above, and its approach—studying behaviour and neural connections in simple animals such as insects (Hoyle, 1984)—concerns the same subject matter. In this thesis, the much-studied behaviour of cricket phonotaxis (how crickets walk towards the calling song of a mate (Huber and Thorson, 1985)) is the particular example of sensorimotor behaviour that will be used.

This system is an interesting one for many reasons. First, it is one where the behavioural consequences of perception are particularly clear, and thus studying

the 'complete' mechanism is possible. It has been the subject of a great deal of behavioural and neurological investigation, so there is a good basis for building a detailed model, and for evaluating the results (Schildberger, 1988). On the other hand, it is perhaps typical of biological examples in that the actual mechanism that underlies the behaviour is not well enough understood that it can be simply programmed into a robot (Schildberger and Horner, 1988). Therefore an investigation through model-building has a potential contribution to make to the understanding of this system, by requiring a complete specification of hypotheses and enabling the plausibility of these hypotheses to be tested.

In Chapter 4, a detailed discussion of this biological system is given. In the process I derive a new hypothesis about the neural mechanism, based on the idea that the processes of *recognizing* a calling song and *walking towards* a sound, which are treated as separable subfunctions in the usual analysis of the cricket's behaviour, may in fact be joint properties of a single, simpler mechanism.

### 1.2.2 A robotic model

A limitation on most symbolic models of sensorimotor processes is that simplification will occur in the critical area of interaction with the environment, and consequently the essence of the perceptual problem may be missed (Brooks, 1986). There is particular value in showing that a mechanism works for real control in the real world, because it reduces the possibility that the 'world' of the model is designed to make the hypothesis succeed—it is sometimes possible to adapt a physical world to suit the mechanism; but, when the world itself is being modeled, it is more likely that only those aspects of it that the mechanism can handle will be represented.

The design and construction of a robot model of the hypothesis about phonotaxis is described in Chapter 5. Using a robot allowed sound to be represented by sound, and movement by movement, and the dynamic temporal relationship between sensing and acting to be preserved. An electronic circuit was devised to capture explicitly the physical capacities of the sensors; and the control program

was based on hypothesised neural processes. Though compromises in details of the model were necessary in several areas, it was still possible to represent the mechanism in sufficient detail to provide an adequate basis for testing the hypothesis.

### 1.2.3 Testing the mechanism

Choosing a particular biological system to model facilitates the evaluation of a model perceptual system; that is, it allows the interpretation of the results to be explicitly related to a target system, rather than merely drawing vague analogies. Similarly the concrete nature of a robotic model is reflected in the results—it is possible to say that a mechanism *does* perform a certain task without requiring justification for such an interpretation.

It is true of AI in general and robotics in particular that evaluation of systems is often cursory (Cohen, 1991). It is sometimes argued that so many factors affect real world behaviour that it is hard to decide on the true effectiveness of a mechanism. However, biology and psychology provide a wide-ranging set of procedures for measuring behaviour: in particular, in the use of experimental procedures and statistics to establish conclusions when there are uncontrollable sources of variation. Adopting such procedures to evaluate the behaviour of the robot was an important step in establishing a useful methodology.

This made it possible to carry out effective tests of whether the hypothesised mechanism embodied in the robot really can explain the phonotactic behaviour of the cricket. In Chapter 6, a series of experiments are described that evaluate whether the mechanism can account for the cricket's ability to find specific sounds. The robot's behaviour was found to be similar to that of the cricket under a variety of measures and conditions. The results are sufficient to establish that the hypothesis is worth further investigation in the cricket.

## 1.3 Outcomes

The work described in this thesis resulted in an effective sensorimotor control mechanism that enables a real robot to find its way to a particular sound source in a realistic environment. While this is not a particularly high level of behavioural competence, it is the kind of basic ability to interact successfully with the world that is said to be lacking in many 'high level' AI systems. Testing shows the mechanism to be reasonably robust and adaptable to a variety of conditions.

Consequently it demonstrates that this mechanism, hypothesised as an explanation of cricket phonotaxis, is sufficient to account for a number of aspects of the cricket's behaviour; including abilities that have been taken as evidence of rather more complex neural processes. Thus the hypothesis, which is simpler than other current explanations, is a plausible one that can act as a basis for study of behaviour and neurophysiology in the cricket. Several directions for further experimentation are suggested by it.

The results support the argument that it is more useful to consider perception in terms of specialised structures that provide appropriate linkages between an agent and its environment than in terms of processes of representation. The hypothesis is based on the idea that the behavioural function of sensory processing is essential to understanding perceptual systems, and that the mechanisms are largely determined by the physical details of the interaction of agent and environment. Such considerations also affected the development of the robot, and yielded an effective mechanism for sensorimotor behaviour.

Most importantly, the work demonstrates a methodology whereby the concerns of biology and AI can be usefully combined, at the level of developing and testing specific hypotheses about sensorimotor mechanisms. It is shown by example that it is possible to usefully investigate perceptual systems by building robot models of biological mechanisms.

# Chapter 2

# Perceptual Systems

## 2.1 Biological Systems

The idea of using simple systems to investigate processes that also occur in more complex systems has always been part of the approach to understanding biological mechanisms. That is, problems general to biology, such as genetic codes or neural chemistry, have been studied using lower animals, such as the fruit-fly or the sea-slug, because their relative simplicity makes it possible to study the complete functioning of the mechanism. They are sometimes referred to as 'model' systems (Rosen, 1983).

Approaching the problem of understanding behaviour in this way has recently seen a resurgence, in the establishment of the field of neuroethology. Camhi (1987) describes the strategy as "selecting a simple, fixed behaviour and then examining from input through output the underlying neural circuit". Improved neurological techniques have made possible the identification of specific neurons involved in particular sensorimotor responses. Clear association of identified neurons with behaviour is currently only possible for animals such as arthropods which have thousands rather than thousands of thousands of neurons (Hoyle, 1984), although some identification of the function and connections of larger neuronal groups has been possible in vertebrates (Ewert *etal*, 1981).

Even with these simple animals, the behaviours investigated are often the simplest stimulus-response reactions such as digestion, fixed action patterns, or escape from noxious stimuli. In other more complex behaviours, such as continuous motor adjustment to a sensory situation, complete neural mapping of the control has

not yet been achieved , although much current work is being directed towards that end.

The degree to which the study of such systems will actually be applicable to other animals, such as higher mammals, is not certain. The very reasons behind the choice of systems to study tend to limit the likelihood of extending the conclusions. For example, Krasne and Wine (1987) say they look at the escape response in crayfish because it has few synapses, a simple circuit and large axons, but admit that because this is a specialised system, it is not good model for generalization. Although Kupferman (1984) may claim that "invertebrates have already provided, and will continue to provide, insights into the neural mechanisms of general classes of behaviour that appear to be common among virtually all animals", Hoyle (1983) criticises the "extremely naive dogma around, based entirely on wishful thinking, that there must be common principles of neuronal organization and function so that these 'simple systems' could serve as models for more complex ones, including mammals". Such systems may nevertheless be useful models for robots, which as yet are no more complex than insects.

While it seems likely that at the level of neural mechanisms, it may be difficult if not impossible to obtain generalizations, the results from these biological studies can still provide useful principles about the nature of perceptual systems, that is, *characteristics* of these systems that seem essential to their successful functioning. In particular, I will discuss here some characteristics that seem to be generally disregarded in current AI approaches to behaviour, despite being strongly indicated by those biological systems that have been closely studied.

### 2.1.1 Specialization

A striking feature of perceptual systems in nature is that their successful functioning often depends upon the fact that they are used for very particular tasks. Such sensory systems don't provide for general interaction with the environment, but only the particular interaction that is relevant to the organism. This greatly influences, and usually simplifies, the mechanism required. This may range from

the occurence of large eyes in nocturnal predators to examples such as that given by Iacino *et al* (1990) "Stomatopods catch their prey with a fast powerful strike of their raptorial appendages ... Eyes of stomatopods differ from all other arthropod eyes".

This feature seems particularly obvious in lower animals, where specialization may occur within a species, for example, the ant caste system, or even within an individual (the developmental stages of insects usually have a phase which is dedicated to feeding, such as the caterpillar stage for moths).

While the simplicity of insect systems may make specialization seem more necessary, examples abound throughout nature, as any natural history study makes clear. Understanding the variety of animal forms requires appreciation of the variety of niches, or task situations, that they occupy. While all animals are concerned to feed, the nature of the food required leads to vastly different means of recognising and obtaining it, and there is little to be gained by equating the caterpiller's choice of leaves to chew with the lion stalking its prey: they require different perceptual systems.

### 2.1.2 No central representation

In any animal systems that have been studied in sufficient depth it seems that the sensory-motor functioning is not mediated by any kind of central 'world-model'. Behaviour is the result of many parallel sensor-actuator connections, not of the sensory situation as a whole being mapped onto a collection of possible actions. Although it is quite common for biologists to adopt a centralised information-processing description of behaviour, the explanatory usefulness of such accounts is doubtful. Jamon (1991) gives a number of examples for animal navigation where assumptions about the need for central maps have proved unnecessary.

Some examples will illustrate what is meant by multiple connections rather than central models. Ferretti and de Talens (1975) describe "several quite distinct visually-controlled behaviour patterns" in the fly, including: optomotor turning in walking and torque in flight; approach to a stationary distant object; escape

from an approaching object; and landing in response to an expanding visual field. Each behaviour uses different cues from the visual situation, and it appears that separate parallel filters operate to detect the relevant cue and initiate the relevant behaviour. Thus optomotor turning depends upon the average relative rotation of the visual field; and the landing response on symmetric motion throughout the field. Similarly, Roberts (1987) describes skin sensory modalities in amphibian embryos, identifying different neural systems underlying the behaviour resulting from i) touch ii) damage or iii) slow pressure, which respectively initiate locomotion, escape or freezing behaviour.

While it is possible to imagine, for these examples, processing that takes in all the sensory information to model the situation and then decides which response is appropriate, the fact is that, at the neural level, this is *not* what actually happens: instead, there are separate processing channels for each behaviour. This arrangement makes sense in evolutionary terms, but it might also be a contributing factor to the success and robustness of the overall behaviour of these organisms.

### 2.1.3 The organism/environment interaction

The morphology of receptors and actuators generally exploit constraints imposed by the task and the environment to simplify the actual neural processing required. This point is related to the first in that the specificity of tasks is a major source of constraints, but here it is emphasised how the physical interaction between the organism and its environment is a critical part of the mechanism for behaviour.

Wehner (1987) presents an excellent discussion of this principle which he calls 'matched filters'. He gives a number of examples, ranging from the optics of crab eyes that exploit the horizon-dominated visual field to obtain size constancy irrespective of distance, to the retinal polarization filter in bees and ants that detects the symmetry plane of the sky, rather than individual directions of polarization. These illustrate the 'trick' of incorporating "the fundamental spatial aspects of a navigational problem into the spatial design of the sensory surface, and then to rely on rather simple circuitry to process the outputs". In other words the

physics of the receptors are an essential part of the mechanism for controlling the behaviour. The physics of the actuators can be similarly critical: control of flight can't be understood without considering the aerodynamics of the wing.

This aspect of natural perceptual devices is worth stressing. Although it seems obvious even under the information processing paradigm of sensory-motor control that the central processing required depends upon the nature of the input and output, the possibility of having the 'smart' processing built into the physics of the interface is often overlooked. Yet this seems to be essential in many natural systems.

### 2.1.4 Actions determine sensations

One motivation for requiring the inclusion of behaviour in the study of perceptual systems is the fact that behaviour critically affects what the sensors encounter. Particular movements are often necessary for efficient sensing. Rasnow *et al* (1988) note that "the active positioning of receptor surfaces directly affects the content and quality of the sensory information received by the nervous system. Thus controlling the position of sensors during sensory exploration constitutes an important feature of an animal's strategy for making sensory discriminations".

The example they examine is the electric fish, which needs to adjust its position to perform various types of discrimination of the objects that distort the electrical field it emanates. Heiligenberg (1987) has also studied in depth the ability of the electric fish to alter the signal it produces to prevent interference when within range of the field of another fish. Many similar cases can be found: for example, it is a feature of many visual systems that there must be retinal movement for objects to be detected at all; reaching out to touch and manipulate an object is often important for identification.

A further aspect of behavioural involvement in perception is the fact that the strategy adopted for the task can reduce the requirements on the sensory mechanism. Most insects have only two long antennae but they can move them around to effectively explore the entire area in front of them. Wehner's example of po-

larization patterns illustrates this also: the insect aligns itself with the symmetry plane of the sky and then turns to set its course, rather than calculating its direction from the polarization pattern directly, which is theoretically possible but computationally more difficult. Another example is the escape response in the lacewing: escape usually requires directional sensing to chose which way to move, but the small size of its ears means they are unlikely to be directional. However, it escapes its main predator, the bat, by folding its wings and doing a 'nose-dive'. This response, requiring minimal processing, greatly reduces the chances of capture (Miller, 1983).

The importance of behaviour in determining sensory input is somewhat overinterpreted by Powers (1973) who contends that behaviour *is* the attempt to get certain perceptual input (generalizing the concept of servo-mechanisms to explain all motor control). The interaction between sensing and action is not a matter of one driving the other, but rather of both being parts of a system that needs to be regarded as a whole to be properly understood.

### 2.1.5 Limited role of learning

Hoyle (1984) states that "For the animal kingdom as a whole, the majority of even the most complex behaviours ...fall into the category of instinctive acts. They require no experience of the behaviour in its context, nor learning, for their perfect execution". This is perhaps an overstatement. It is becoming clear that many 'instinctive' behaviours can be modified somewhat by environmental contingencies, and may require environmental tuning for 'perfect' execution. On the other hand, it also appears that the extent to which behaviour can be modified, even in higher animals, is quite strongly constrained by genetic dispositions. A clear division between completely innate and completely learned acts is not possible.

Habituation and sensitization effects are often included as examples of 'learning' in lower animals (Eisenstein and Reep, 1985) but this seems an unnecessary conflation of two kinds of modifiability of behaviour. It seems wrong to imply that habituation is not innate, as the effect is a predictable result of the genotype of the

organism. Learning involves more long-term alteration of behaviour that requires reference to the organism's experiential history to explain. Again the distinction is not absolute, 'long-term' and 'history' are not fixed quantities, but there seems a qualitative difference between a temporarily reduced reaction to rapidly repeated stimuli, such as the gill withdrawal reflex of *Aplysia*, and the ability to relocate a food source using a learned route, as exemplified in the honeybee.

A large proportion of behaviour in lower animals does appear to involve hard-wired perceptual mechanisms: the fact that single neurons can be identified as having the same morphology and function in each individual of a species is a strong indication of this. This seems particularly true of the most basic functions, such as controlling digestion, maintaining stability while moving, and mating behaviours. Moreover, such fixed functions are better understood— Schwartz (1988) suggests that "theorists who take some hypothesis about learning as their starting point are choosing to begin in a particularly dark area of neuroscience". "The generally modest level and extent of insect learning" (Hansell, 1985) does not prevent them from performing successful and robust behaviour in a complex and dynamic world.

Of course, these hard-wired systems are the result of evolutionary adaptation. However it is important to avoid the fallacy of conceiving of this process as one of design, in that there is no non-recursive way to evaluate the 'success' of the design. Pierce and Ollanson (1987) point out that at every stage of its evolution, a sensory system has to be functional, and it is always possible to retrospectively define a task and situation for which it is ideally adapted, but "the very process of evolutionary change constantly redefines the material available to work with and the problem to be solved". This has consequences both for attempts to understand perceptual systems by abstract analysis of the problem (such as optimal foraging theories), and for attempts to build perceptual devices by 'evolutionary' processes, which will be discussed in the next chapter.

### 2.1.6 Complexity

A final point that emerges from consideration of the work that has been done with these 'simple' systems is the startling degree of complexity that is revealed by close examination. No complete neural map of a sensorimotor connection has yet been achieved: though there is progress in that direction for a few systems it has required an enormous amount of investigation to get that far. Further, the more investigation advances, the less common ground there appears to be: behaviour seems to be the result of a variety of special mechanisms rather than the application of a few general laws. "There is no *generally* applicable description of the forces underlying 'animal behaviour'. Rather the enormous complexity and diversity of the behaviour of animals requires specific models for the analysis of specific issues" (Fernald, 1984).

An example is orthopteran song production (Ewing, 1989): cicadas, grasshoppers and crickets are closely related species and have the same 'task' of producing a loud regular sound to attract mates. Cicadas have a collapsible cuticle area on their abdomen backed by resonant air-sacs; grasshoppers draw a leg file across the forewings which radiate the sound; crickets move one wing across the other which has a resonant surface area; some cricket species dig burrows that improve the acoustic efficiency. Understanding one of these sound production systems does not greatly contribute to understanding the others. Hoyle (1983) poses the question "The starting point of all behaviour must be the activity of neurons, but is it an emergent property, resulting from the interaction of many having like properties, as it at first seemed, or might it reside in unique individual properties of components of specific circuits?" Neuroethological research so far seems to indicate the latter.

### 2.1.7 Summary

A number of interesting characteristics emerge when biological perceptual systems are examined. Such systems are specialized, and utilize specific aspects of their physical interaction with the world to achieve robust behaviour, with minimal

central processing or learning. All these aspects will be further illustrated in Chapter 4 when phonotaxis in the cricket is examined in depth.

## 2.2 Artificial Systems

A different perspective on perceptual systems is provided by the attempt to build new ones, rather than take existing ones apart. Finding the sensory-motor linkages required for successful behaviour in a robotic system is in essence the same problem as that solved by the organisms explored by neuroethologists. This particularly applies to robots that can get about on their own in the real world: "Working on autonomously guided vehicles is clearly tantamount to working on a kind of holistic animal microworld: such work is forced to respect many (but not all) of the constraints that we saw would apply to evolved biological systems" (Clark, 1989).

### 2.2.1 Robotics and AI

For most of its history, the study of Artificial Intelligence has been dominated by attempts to get computers to do intellectual tasks. Mathematical reasoning, logic, problem solving and game playing all proved possible to implement 'artificially', and it appeared possible that computers could be made to think. However, progress towards complete intelligence slowed down when it came to getting computers to extract and formulate the problems to be solved from natural language descriptions or real-world situations.

There was always a substrand of attempts to build systems that could operate competently in the real world—that could exhibit behavioural, rather than intellectual, competence. Sometimes this was explicitly linked to animal behaviour, especially in the cybernetic tradition. However the approach came to be influenced by 'intellectual' AI, in that the problem of getting behaviour was conceived as a matter of reasoning from sensor inputs to motor outputs, analogous to the

reasoning from premises to conclusions of problem solving, or the choice of a move based on the current situation of game playing. The resulting 'sense-think-act' robots (Malcolm *et al*, 1989) would take in all sensor readings, process them to get an answer, and then output the move.

Robots such as *Shakey* (Nilsson, 1984) or the Stanford cart (Moravec, 1983) are examples of this approach, neither of which produced particularly successful or interesting behaviour (Moravec's cart ran at 3 to 5 metres per hour, was designed to reach a target location while avoiding obstacles, but hit obstacles on three of five runs, and failed to finish a 20m course due to decreasing battery power on the other two). Formidable problems were raised in trying to get the sensory information into a form suitable for thinking about, and linking the results of these thoughts to actual motor control. This led to a subdivision of research efforts to tackle each part of the problem (sensing, thinking and acting) on its own. To do the thinking on its own required a 'representation' of the external situation so that there was some form of input to be transformed to some form of output.

This division of forces and reliance on representation is the current mainstream. Separate research groups work on the sensory problem of transforming sensor signals into a usable world model (particularly in computer vision) and the motor problem of translating a model plan into motor execution (particularly in industrial robotics). Meanwhile the work being done on 'thinking' makes many assumptions about what sensory and motor competence might eventually be available, and goes on to devise control systems on that basis. Several of the 'classic' problems in AI are due to working on the central control on its own: it has led to the development of systems that have no direct link to the world which produces the problem of 'symbol grounding' (Harnad, 1990); it involves the attempt to represent internally the knowledge that the system requires (as a world model) which leads to the 'frame problem' (Pylyshyn, 1987). As the system is no longer being studied as a whole, it has also affected the understanding of the nature of perception. For example, Hurlbert and Poggio (1988) ask "What does vision do? The plain answer is that vision transforms light signals into internal representations of the things

that transmit them". This may be true of current computer vision systems, but it does not suffice to explain the purpose of vision in biological systems.

### 2.2.2 The complete system approach

There remained an undercurrent of research committed to the idea that the task of getting a complete complex system was best approached by building complete simple systems rather than partial complex ones. One of the earliest mobile robots, the 'turtle' of Grey Walter (1953) was such a system. Recently there has been something of a revival of this approach, particularly by Brooks and his colleagues at MIT. It is convenient to first discuss their observations about what seems to be fundamental in the design of successful simple robots, as they have explicitly rejected the traditional approach in favour of finding out from experience 'what works'. This has resulted in ideas about perceptual systems that will be further explicated in the following sections.

First, MIT's successful robots have resulted from setting out to solve particular problems: getting certain behaviour from a certain robot in a certain environment, rather than trying to solve some general, abstract problem of behaviour. They have found that considering sensory processing in terms of the behaviour it is supporting tends to enormously simplify the problems usually posed in interpreting real sense data. For example most of the robots move around in the world without any idea of 'where they are' because they don't need this information to avoid obstacles and find what they are looking for, such as soda cans (*Herbert*), or another robot to follow (*Tom and Jerry*), both described in Brooks (1991b). Indeed, as pointed out by Horswill and Brooks (1988) even obstacle avoidance isn't needed by a chasing robot as it can assume the target it is following hasn't gone through an obstacle.

Having specific tasks means that the processing is aimed at connecting sensing to acting, with as few intervening steps as possible. A result is that internal representation, in terms of a model of the world, can generally be abandoned. Brooks (1991a) argues that this has an additional advantage "If there are no models built, the problem of uncertainty is inherently reduced". The alternative "is to operate

in a tight coupling with the world through a sensing-acting feedback loop". For example, *Genghis* "directly implements walking through many very tight couplings of sensors to actuators" (Brooks, 1991b). *Herbert* (*ibid.*) coordinates the various stages of its task by having each subtask detect when the correct environmental conditions for execution exist. The results seem to indicate that abandoning representation can be an effective strategy in getting working systems.

Another observation that has emerged from this work is that "perception and motor skills are the hard problems solved by intelligent systems, and ... the shape of solutions to these problems constrains greatly the correct solutions of the smaller pieces of intelligence that remain" (Brooks, 1991c). This notion of sensory and motor mechanisms constraining central processes seems to arise whenever a system is embodied. "If the intelligent agent has a body, has sensors, and has actuators, then all the details and issues of being in the world must be faced ... it is usual to find that many of the problems that seemed significant are not so in the physical system (typically 'puzzle' like situations where symbolic reasoning seemed necessary tend not to arise in embodied systems) and many that seemed non-problems become major hurdles (typically those that concern aspects of perception and action)" (Brooks, 1991d). For example, in implementing robots, Flynn and Brooks (1989) report that more time was spent battling noise in infrared receptors than in writing subsumption code:"We have found that perception is the key problem while intelligent control is relatively easy in comparison".

The same paper also introduces the slogan "let the physics do the walking". The physics of transducers, if exploited, can lead to simple and robust solutions to problems of behaving, often involving very little processing. Having the *right kind* of physical abilities for the task may be much more useful than having particularly accurate or powerful sensors and actuators, or a large processing capability.

The general picture emerging from this research can be summarised as follows:
• The internal mechanisms should be approached in terms of the specific behavioural outcomes.
• This is best accomplished by direct coupling of sensing to action (through the robot) and action to sensing (through the world): there is no role for intermediate

central representations.

- The functioning of the sensory and motor mechanisms determine the control mechanisms.
- Having the right physics and peripheral processes in the sensors and actuators can be a major factor in getting the right behaviour.

It is also worth noting that most of the MIT robots have had hard-wired control circuits, rather than circuits developed by learning or evolutionary processes, although some parts of the systems have been subsequently reimplemented using such mechanisms (Maes and Brooks, 1990).

It is not possible within the scope of this thesis to review the entire field of robot research to evaluate the level of support for these ideas, but some further elaboration and examples of each of these points is given in the following sections. In addition to Brooks, others with wide experience of the field have raised similar points (Malcolm *et al*, 1989; Moravec, 1984), although not all research on simple autonomous systems reflects these conclusions: the mobile robot problem is still often discussed in 'traditional' terms. For example, despite agreeing with the insights of Brooks in their introduction, Henderson and Grupen (1990) go on to say that "Logical behaviours for obstacle avoidance" *must* include "incrementally build a 3-D representation of the world (i.e. determine its motion and integrate distinct views into a coherent global view) ... explicitly represent, manipulate and combine uncertainty, and build a semantic representation of the world". It seems, however, that in a significant proportion of actual, working, robotic systems, as opposed to theoretical robot architectures or computer simulated systems, the principles above have application.

### 2.2.3 Mechanisms determined by behaviour required

Any designer of any system is bound to be aware that the task that the system is to perform specifies what capacities the mechanisms making up that system must have. Abstract ideas about what capacities are necessary for intelligent action have rarely turned out to be the same as the specific requirements for actual sensory-

motor tasks. Many successful robots work by utilizing the special assumptions that can be made when a particular behavioural context is specified. A simple example is that, contrary to Henderson and Grupen's contention above, it is not necessary to compute locations and identities of objects in order to get around them.

Arkin's (1989, 1991) schema-based robot navigation adopts this principle: "Most significantly, perception should be viewed as action-orientated. There is no need to process all available sensory data, only that data that is pertinent to the task at hand." This allows substantial reductions in computational demand by sensory processes, and behaviours such as avoiding obstacles, staying on paths and moving to goals can be implemented by having sensing specific to each task, such as ultrasonic for obstacles, fast region segmentation on an image for path following, and shaft encoders for approximate goal location.

Another aspect of the same idea is that particular ways of behaving can be used to alter sensing requirements. Recent work in the field of animate vision is based on the idea that looking at vision in context of system's behaviour radically changes the nature of the problems, and frequently reduces or even eliminates them. Ballard (1991), discussing gaze control mechanisms, notes that when "visual calculations are embedded in a sensory-motor behavioural repertoire" those calculations can often be greatly simplified.

### 2.2.4 Behaviour through coupling not central representations

The idea of using multiple sensor-actuator links, rather than having all sensory information contribute to a central representation that is then interpreted to decide on an action, has been successfully adopted in several robot systems in addition to those described by Brooks. Arkin's work, mentioned above, makes use of this approach, in which "sensor data is normally channeled directly to the individual subtasks" that in combination produce the required behaviour.

Payton and colleagues use a similar approach to obtain autonomous vehicle control, with an architecture "designed specifically to handle diverse terrain with maximal speed, efficiency and versatility through the use of a library of reflexive strategies specialized to particular needs" (Payton, 1986). The lowest level of control uses virtual sensors and reflexive behaviours, the former detecting specialized environmental features and the latter utilizing a virtual sensor to generate real-time control. Designing reflexive behaviours makes use of assumptions about the environment to allow rapid processing. The resulting system is capable of cross-country autonomous operation (Daily *et al*, 1988).

Anderson and Donath (1990) similarly use "a set of *primitive reflexive behaviours*, each of which causes the robot to respond to a subset of the total stimuli". Using combinations of several avoidance and attraction behaviours they demonstrate more complex behaviours emerging from the interaction, such as simple navigational competence. The system works robustly and in real-time on a robot.

Many of these researchers suggest that central representation might ultimately have some role in obtaining more complex behaviour such as building maps or carrying out plans. However this would be in addition to fundamental 'reactive' abilities involving direct connections, not a substitute for them.

### 2.2.5 Effects of sensing and motor mechanisms on control

Many robot researchers note the degree to which controlling the robot is determined by exactly how the sensory and motor processes work. This is revealed in the fact that a great deal of robotics research effort is curently aimed at producing more accurate sensors and actuators. 'Accuracy' is required by the traditional architecture, but the alternative approach described here suggests that the effectiveness of various kinds of sensor and actuator can only be assessed relative to the functioning of the complete system. It is quite possible for sensors to be *too* accurate for a particular use, necessitating more complex processing to extract a simple environmental property. For example, a robot that cannot perceive gaps

smaller than its own width doesn't need additional mechanisms to reject such gaps when performing obstacle avoidance.

'General' systems for robotic control have been difficult to devise because working systems are highly dependent on the details of the mechanism in which they are implemented. Researchers working on large-scale automous control projects such as *Ambler* (Whittaker *et al*, 1991) or *NavLab* (Singh *et al*, 1991) have noted that developing systems in simulation that will work when implemented in the robot is critically dependent on having accurate representation of the sensory and motor interface to the environment (this point will be pursued in the next chapter).

For researchers working on complete behaviours in simple systems, the importance of the sensing and motor mechanisms is unavoidable. Nehmzow and Smithers (1992) quote Clark (1987) "both the perceptual and motor capacities of the system in which implementation occurs are crucial" and note "Our work on building real robots that work reliably in real environments confirms this insight". Pichon *et al* (1989) similarly note that "Attempts to design this reflex-like visuo-motor system emphasised the tight interdependency between sensory and motor processing, each of which sets it own constraints on the other".

### 2.2.6 Getting results by having the right physics

Another way of looking at the previous point is that manipulation of the physical characteristics of sensors and actuators can greatly simplify the control strategy. The interaction between the robot and the environment at the physical level can facilitate the particular task to be performed enormously. A simple example is that robots of certain shapes may be less likely to get stuck in office enviroments.

Malcolm (1990) mentions the potential application of this principle in assembly robotics. "It is possible to make use of prior knowledge about the assembly task to simplify the sensing task very considerably. It is even possible in some cases to take advantage of naturally occuring constraints on motion to provide the feedback to control errors in part motion without requiring sensing at all, such as by pushing or dropping actions".

Recognition of the physical constraints imposed by particular task environments has also been explicitly adopted as a strategy for mobile robot control. Horswill (1992) proposes the identification of "habitat constraints" which describe "conditions on the structures and processes of a creature's environment" as a means of facilitating computational problems. For example, he uses the assumption that objects rest on a flat plane to judge their relative distance from the robot by having a camera that points forward and somewhat downward, so more distant objects will be higher up the image.

Raibert's (1989) work on hopping and running mechanisms similarly exploits physical dynamics to reduce the complexity of internal control. For example he uses the technique of symmetry: "Symmetry simplifies the control because it frees the control system from regulating the details of the trajectory—the details are determined passively by the mechanical system" (Raibert, 1986). Koditschek and Buhler (1991) sum up this approach as "manipulation strategies that achieve desired goals by systematically harnessing the intrinsic features of the robot-environment interaction". The success of mechanisms designed in this way indicate that it is a powerful strategy.

### 2.2.7 Conclusions

These ideas emerging from building artificial perceptual systems obviously have links to those discussed for biological systems. Together, these observations suggest that explaining perceptual systems will involve thorough understanding of the interaction between the system and its environment: in terms of the behaviours of the system; the physics of the interface; and how the system connects sensors and actuators through itself and through the world.

Such an approach does not fit easily into the now traditional characterisation of perception as information processing involving manipulation of representations. In this view, perception involves the extraction of information by the organism from the external world. The information to be extracted is often characterised as general knowledge rather than specifically related to the ongoing behaviours of

the agent; when it is so specified, it is generally assumed that supporting those behaviours requires explicit internal representation of the relevant information. Either way, the emphasis is placed on internal processes to produce representations. By contrast, the emphasis here has been placed on the *interaction of the internal and external forces that produce behaviour.*

## 2.3 Philosophical Background

There is some tendency to justify the 'simple systems' approach in AI by redefining 'intelligence' as adaptability, which conveniently brings the study of such systems within the field. For example, Beer (1990) writes "What does it mean to be intelligent? ... I would like to argue that it is *adaptive behaviour*, the much broader ability to cope with the complex, dynamic, unpredictable world in which we live, that is, in fact, fundamental"; Winograd and Flores (1986) say "The essence of intelligence is to act appropriately when there is no simple predefinition of the problem or the space of states in which to search for a solution". Although it is often taken as a corollary, the idea that such study will be of significance to understanding traditional intelligent functions is not supported by this redefinition alone.

The most common arguments given in support of the contention that work on simple sensory-motor behaviours will scale up to human intelligence are: that rational thought appeared very recently in evolution and was probably dependent on the development of sensory-motor competence that went before; and that much of human behaviour is not determined by rational thought, but rather by unconscious competence of the kind displayed by animals. Philosophical views of behaviour and intelligence are often used to support these ideas.

However, the validity of the approach to understanding perception in animals and machines that has emerged from the previous sections does not depend upon proof that the ideas will explain human intellect: understanding perceptual systems is an interesting pursuit in its own right. Rather, the reason for exploring

some philosophical views here is to provide additional support for the argument, so far based on practical examples, that the 'processing of representations' explanation of perceptual systems may not be the best one.

### 2.3.1 Direct perception

The view of perception that has been presented here has much in common with Gibson's (1979) 'ecological' approach to perception. In (Gibson, 1961) he stated that "The overall problem of vision is that of understanding those activities of men and animals which depend on the stimulation of their eyes", and throughout his work he stresses the importance of understanding perception in terms of the activities of the perceiving system. Thus he claims that *what* is perceived are 'affordances', *i.e.* the possibilities for action that the environment affords, and that these lie in invariances in the flux of sensation. For example the *rate* of expansion in the visual field affords the time to contact with an approaching object (Lee, 1980).

It is worth pointing out that the use of 'perceptual system' in this thesis is a little different to Gibson's use of the same term. Gibson (1966) distinguishes between exploratory movements, which are part of the perceptual system that they move (thus retinal focusing, eye orientation and head and body movements that change the position of the eyes are all part of the vision system) and performatory movements that may be guided by the 'information' provided by the perceptual systems, but are not part of the system itself. In his examples, however, this distinction is not always clear (for example, in orienting behaviour).

'Information' here refers to the 'affordances' of objects, *i.e.* information is meaning. In Gibson's view the meaning is not constructed in processes of perception, rather, perceptual systems are structured, through development as well as genetically, so as to register directly the meaning that is inherent in the environment. This notion of 'direct perception' has been criticised as denying the role of internal processing as relevant to understanding perception (Ullman, 1980), but rather it denies the validity of the concept of internal representation as an expla-

nation of perception. It is true, however, that Gibson's notion of direct pickup through neural 'resonance' is never elaborated into explanation of *how* this takes place, possibly because he does not consider the details of neural processes to be part of a psychological explanation.

Although most current perceptual theory is based on the information processing view, the possibility of an alternative has been well expressed by Braddick (1980) "... it has not been established that information about a stimulus variable has to be explicit in any particular, obvious sense at some level of representation in order that this variable should be available to concious perception or that it should drive verbal or motor behaviour". While it does seem worthwile to pursue explanatory mechanisms for perception, which Gibson avoids, these need not be the representational mechanisms currently popular, and Gibson's *description* of perception provides an alternative way to approach the problem.

### 2.3.2 Against representationalism

Conceptual difficulties with a representational account of perception have been pointed out by the realist tradition in philosophy. Maze (1981) examines the difficulties of a coherent account of the function of mental representations as mediators in perception. If we can only know directly the mental entities that act as representions of the world, how can we know what they represent? The relationship of representation requires knowing both terms: for A to stand in for B (for some observer S), S must know both A and B. But if we *do* already know the things they represent, then the representations no longer have a function as mediators of what we know. Unlike Maze, I would argue that they may still have some function as a means of manipulating what we know; a mental representation is equivalent to a physical representation, such as a diagram, which is a tool for reasoning (Clancey (1992) presents some similar ideas).

The usage of 'representation' in AI is somewhat different to the traditional one in philosophy, which leads to some confusion in criticism. One traditional criticism is that the notion of representations mediating perception simply pushes

the problem back a level as it is then necessary to explain how a homunculus perceives the representations. An attempt in AI to get around this considers the perceiving to be done by *levels* of homunculi, each simpler than the last, bottoming out in homunculi no more complex than neurons (Dennett, 1979). However, such 'self-understanding' representations are then representations only from the point of view of someone observing the process; it is not the fact that they refer to external states of affairs that causes their function in the system (as required by the traditional notion, from Brentano (1874), that mental states are inherently intentional, *i.e.* about something). As far as the system is concerned, the representations (or symbols) are 'opaque' (Fodor, 1981), that is their 'meaning' cannot be accessed, and processing depends entirely upon their form.

This idea of 'meaningless symbol processing' is rather misleading, because 'meaning something' is the only way to identify something *as* a symbol. Harnad (1990) considers the major flaw in symbol processing systems in AI to be the fact that 'meaning' is only provided by the programmer's interpretation; and suggests that the symbols needed to be 'grounded', *i.e.* made meaningful, to the system itself by connecting the computer to the world with sensors and actuators. In other words, the system needs direct contact with the world in order to use symbols to represent the world. But then using symbols to represent the world cannot be the means by which the contact with the world is established: perception must be prior to representation, not dependent on it.

An alternative interpretation is that having internal structures that, to an external observer, are analogous to the external world (*i.e.* represent it), may be the means by which a system performs successfully. For example, a calculator works because its internal wiring is analogous to mathematical operations (Cummins, 1989). It is important to realise that this is an *empirical* hypothesis, rather than a philosophical assumption, and as such requires independent evidence for each system it proposes to explain. As a hypothesis about perceptual systems there is an inherent problem however: the purpose of perception is to transform external signals into something different—concious experience, and/or behaviour—*not* to reproduce them internally.

### 2.3.3 Behaviourism

Behaviourism as a psychological theory and methodology has a number of limitations that have led to its general rejection. Nevertheless, a number of concerns expressed by behaviourists, originally about introspective psychology, more recently about cognitivism, echo ideas discussed above. This is not surprising when it is considered that the subject matter of behaviourist research—how "organisms ... adjust themselves to their environment by means of hereditary and habit equipments" (Watson, 1913)—is a comparable attempt to study simple behavioural systems rather than investigating human conciousness.

Rather than completely rejecting mental states as relevant to behaviour, as is generally supposed, behaviourists challenged the role of such states as *mediators* of behaviour. Ryle (1949) interprets mental ascriptions as *dispositions* to behave, rather than as internal causes of external behaviour. "Overt intelligent performances are not clues to the workings of minds; they are those workings". Ryle also questions the significance of rational intellectual operations in the majority of human actions, and rejects the 'intellectualist doctrine' (which seems very like traditional AI) that defines intelligence in terms of ability to theorise and attain knowledge of truths. He stresses that "In ordinary life ... we are much more concerned with people's competences than with their cognitive repetoires". Consequently he rejects the extension of 'intellectual processing' to explain all behaviour.

Skinner presents some objections to the use of 'representations' as explanations of cognition. The point made at the end of the previous section is put eloquently by him, arguing that sensing must have a behavioural function "At some point the organism must do more than create duplicates ... seeing, hearing, smelling and so on must be forms of action rather than of reproduction"(1984). He argues, for example, that the 'association' of lemons and bitterness doesn't require connecting representations of 'lemon' and 'bitter' in the head: lemons and bitterness are associated in the real world (1977). His comment that "It is most convenient for both organism and physiologist if the external world is never copied—if the world

we know is simply the world around us"(1984) is comparable to Brooks suggestion that 'the world is its own best model'.

However, like Gibson's ecological approach, behaviourism in general chose not to investigate the actual mechanisms in the organism by which sensory 'stimuli' are linked to motor 'responses', focusing on *identifying* regularities in this input-output relationship, rather than *explaining* them (the so-called 'black-box' approach). Again, this has led to criticism, and the now general dominance of the information-processing approach which "promises an account of such relationships *plus* an account of the mechanisms which permit the extraction" (Braddick, 1980). Studying simple perceptual systems in biology and robotics provides the possibility of explanation in terms of neural-level functioning, thus providing mechanisms to flesh out the philosophical objections to conceiving of perception as 'information-processing'.

### 2.3.4 Enactive perception

Building on the concept that the essence of living systems is that they are self-producing ('autopoetic'), Maturana and Varela (1980) regard behaviour as the structural coupling of the agent to its environment, requiring that "the question 'How does the organism obtain information about its environment?' be changed to 'How does it happen that the organism has the structure that permits it to operate adequately in the medium in which it exists?' ". Consequently, they propose "The anatomical and functional organization of the nervous system secures the synthesis of behaviour, not a represention of the world".

Recently, Varela has articulated an approach to perception that has its basis in this autopoetic view of behaving systems. This approach has strong similarities to what has been advocated here (although the characterisation of biological and robotic perceptual systems I have made above was arrived at independently).

Termed the 'enactive' view, it is based on the notion that "perception consists in perceptually guided action ...

> the reference point for understanding perception is no longer a pregiven, perceiver independent world but rather the sensorimotor structure of the perceiver (the way in which the nervous system links sensory and motor surfaces). This structure—the manner in which the perceiver is embodied—rather than some pregiven world determines how the perceiver can act and be modulated by environmental events. Thus the overall concern of an enactive approach to perception is not to determine how some perceiver-independent world is to be recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world." (Varela *et al*, 1991)

In Thompson *et al* (1992) an extended application of this approach to the problem of colour vision is provided; arguing against both the 'objective' view (colours are properties of the world that the perceiver recovers) and the 'subjective' view (colours are constructs of the perceiver projected onto the world). Rather, they argue, colour is specified in the interaction of the world and the perceiver: "Colour can be understood only in relation to the visual perception of a given individual or species (contrary to objectivism); but such visual experience can be understood only in the context of its ecological embodiment (contrary to subjectivism)".

Several other aspects of the discussion in Varela *et al* (1991) echo ideas raised in the previous sections. They reject the use of operations on representations as explanatory for cognition. They also stress the need for a better understanding of the process of evolution, making the point that evolution produces *viable* rather than *optimal* systems. It is interesting that they also use Brooks' research in robotics as an illustrative example of the enactive approach in practice.

### 2.3.5 Conclusions

This philosophical background helps establish a basis for what we need to understand about perceptual mechanisms. Varela *et al* (1991) also discuss the approach required:

> "Consider that there are two domains in which we can describe any cognitive system: on the one hand, we can focus on the structure of the system by describing it as composed of various subsystems, etc., and on the other hand, we can focus on behavioural interactions of the system by describing it as a unity capable of various forms of coupling. In switching back and forth between these two kinds of description, we—that is, cognitive scientists—must determine both how the environment constrains the system and how these constraints themselves are specified by the sensorimotor structure of the system ... In so doing, we are able to explain how regularities—sensorimotor and environmental—emerge from structural coupling. The research task in cognitive science is to make transparent the mechanisms by which such coupling actually unfolds and thereby how specific regularities arise."

The next chapter attempts to specify the role of AI in "making transparent the mechanisms by which such coupling actually unfolds".

# Chapter 3

# AI and Biology

## 3.1 Autonomous agents

The concern with understanding perceptual systems by looking at simple examples of sensory control of behaviour is a growing field in AI. It is frequently described as the study of 'autonomous agents': it is sometimes identified with 'Artificial Life', but that term is also used more widely (cf. Langton, 1989). 'Autonomy' has also been variously defined: for example by Bourgine and Varela (1992) "Autonomy in this context refers to their basic and fundamental capacity to *be*, to assert their existance and to bring forth a world that is significant and pertinent without being digested in advance"; by Smithers (1992) "An agent is autonomous if it is able to cope with all the consequences of its actions to which it is subjected while remaining viable as a task achieving agent in the world it operates in". Definitions always include sensory control of action as at least part of the problem to be solved. Discussions of autonomy usually refer explicitly to simple biological creatures as instances of 'autonomous agents'.

So what is the potential contribution of studying these systems within AI, rather than by traditional biological methods? It is sometimes argued that a more 'general solution' to the problem of *possible* autonomous agents may be found by not limiting study to the *particular* cases found in biology. Such claims seem somewhat specious: surely trying to understand existing autonomous systems is more likely to be informative than undirected attempts to create them? At any rate, the research of interest here is that which claims concern with understanding biological systems, through the use of AI methodology.

Given such concerns, AI's role can be best be characterised as modeling: that is, AI is performing the role in investigating perceptual systems that modeling fulfills in science in general. In this chapter I will analyse this role; justify such an interpretation of AI research; explicate the consequences of this interpretation for evaluation of that research; and draw on this and the observations in the previous chapter to derive some methodological suggestions for how such research might be made to serve the goal of understanding perception.

## 3.2 Models

### 3.2.1 The relationship of modeling

There have been few attempts within AI to analyse carefully the meaning of the term 'model' despite its widespread use (two exceptions are Rothenberg (1989) and Chan and Tidwell (1992); a review of earlier discussions in AI of modeling and simulation is given in Webb (1987)). One source of confusion about the term is how it related to theories and simulations. 'Simulation' is often taken to be synonomous with 'modeling', and 'model' and 'theory' are often used interchangably, yet theories and simulations aren't quite the same thing. It seems best to distinguish two uses of 'model': a 'theoretical model' that is an explanation in some form of language that *describes* the capacities and structures thought to cause a system to exhibit a behaviour; and a 'working model' (or simulation), which is a *physical embodiment* of a theoretical model that acts dynamically to produce data.

Under the latter meaning, a model is an example of the relationship of analogy. An analogy is a three term relationship in which something S takes an object A* to stand in for an object A because either

a) A* and A share certain properties (substantial analogy) or

b) there exists a correspondence between the properties of A* and A (formal analogy). By 'correspondence' is meant a mapping between the properties, which may be quite arbitrary and incomplete.

It is sometimes forgotten that this relationship is a three-part one: after all anything can be incompletely mapped onto anything else, and there seems little point defining a relationship that holds between every pair of objects. There must be some S that makes use of this mapping, i.e. uses A* to stand in for A for some purpose. Chan and Tidwell, for example, have an implicit S in their notion of the 'target type' for a model, which is defined as the 'relevant aspects' of A to which A* corresponds. Rothenberg defines modeling as "the cost effective use of something in place of something else for some cognitive purpose".

Analogy in everyday usage tends to imply pointing out correspondences between two already existing systems. In modeling the correspondence is, for the most part, constructed: that is, the model is *given* components and structure that are meant to correspond to the components and structure of the system, so that the behaviour of the model will correspond to the behaviour of the system.

### 3.2.2 Modeling and explanation

In AI there are two intertwined purposes: to try to explain certain behaviours (the 'scientific' purpose); and to build machines that produce certain behaviours (the 'engineering' purpose). In either case there can be said to be a 'system of interest' $A$: in science it is the existing system producing behaviour $B$ that we want to explain (for example, an insect's avoidance response); in engineering it is a system that could be built to produce behaviour $B$ (for example, an obstacle avoiding robot). In either case, the investigator will hypothesize that the system $A$ has (or will have) certain component properties and structure that cause (or would cause) the behaviour. That is, they will hypothesize a system, $A_H$, with known component properties and structure, and try to show that $A_H$ is the same as $A$, the system of interest, and thereby explain (or produce) its behaviour $B$.

So what purpose can modeling play in this endeavour? To provide support for the conclusion that $A_H = A$, the investigator can, in the scientific case at least, take the actual system apart to look at the components and structure. Alternatively,

or in addition, they can try to show that the proposed system $A_H$ does produce behaviour $B$. That is, they wish to follow the reasoning process

*if* (i) $A_H$ produces certain behaviour $B_H$

*and if* (ii) $B_H$ cannot be distinguished from the behaviour $B$ of $A$

*then it is possible that* (iii) $A_H$ is the same as $A$.

In science this is not sufficient to *prove* $A = A_H$, because there may be more than one way to obtain the behaviour; in engineering, barring further constraints on $A$, producing the behaviour fulfills the requirements.

Discovering (i), *i.e.* what behaviour $B_H$ is produced by $A_H$ may be done by any of several methods. It might be possible to go from the description of $A_H$ to logically find $B_H$. Alternatively, it might be possible to physically construct $A_H$ and observe its behaviour. A third possibility is to construct a dynamic representation $A_H^*$ (*i.e.* a model) and observe *its* behaviour. Modeling is also sometimes considered a development of the first method, as it constitutes a means of finding the logical consequences of $A_H$, particularly in cases where the theory has no mathematical solution.

Thus in modeling, the investigator is constructing a representation $A_H^*$ of the hypothesised system $A_H$. This introduces new steps in the chain of reasoning, as follows

*if* (i*) $A_H^*$ is a valid representation of $A_H$

*and if* (ii*) $A_H^*$ produces $B_H^*$

*and if* (iii*) $B_H^*$ is validly interpretable as $B_H$

*then* (i) $A_H$ produces behaviour $B_H$

*and if* (ii) $B_H$ cannot be distinguished from $B$

*then it is possible that* (iii) $A_H$ is the same as $A$.

The reasons for modeling are well summed up by Naylor *et al* (1966): "The model is amenable to manipulations which would be impossible, too expensive or impractical to perform on the entity it portrays. The operation of the model can

be studied and, from it, properties concerning the behaviour of the actual system or its subsystem can be inferred".

Determining the strength of the evidence provided by the result of modeling is dependent on determining the accuracy of the representation (i*) as well as the production of the behaviour (ii*) the interpretation of the behaviour (iii*) and the comparison of that behaviour with the actual system (ii). In other words, reasoning about a system on the basis of a model of that system is constrained entirely by the mapping between the system and the model. While this may seem logically obvious, failure to appreciate this constraint is endemic in science. Michel (1986) provides an excellent discussion of this problem in relation to the use of measurement in psychology, and Rosen (1983) applies similar arguments regarding biomedical investigations.

### 3.2.3 Consequences for AI systems

It seems fairly clear from the above characterisation of the role of modeling that AI *is* modeling: AI researchers build physical systems that correspond to hypothesised systems that explain certain behaviours, and they examine the behaviour of the the resulting systems to see if it corresponds to the behaviours of interest as a means of evaluating the hypotheses. Rothenberg says "AI ... views the implementation of computerised models as a key technique for understanding intelligence" (1989); Clancey (1991) likewise suggests "We can generalize what AI programming is in terms of a modeling methodology".

This role for AI seems particularly clear when discussing the investigation of perceptual systems for autonomous behaviour. The common concern is to find the mechanisms that cause the behaviour. In neuroscience, the plausibility of hypothesised mechanisms is explored by looking for the existence of those mechanisms in actual biological systems. In ethology, aspects of biological behaviour are used as evidence to support the likelihood of certain underlying mechanisms. AI can be seen as an extension of the normal role of theorising in this field: that is, AI provides a means of suggesting, making explicit, and exploring the consequences of

hypotheses about mechanisms, which can then be used to guide further research; and it does this by building models.

The status of robotics research in this area needs to be understood carefully. Insofar as 'autonomous behaviour' is the aim, a robot may constitute the building of a hypothesized system; and indeed a 'robot' is often the system of interest that is being modeled using purely computational representations. On the other hand, if the aim is to understand biological mechanisms, a robot is then a *model* of the hypothesized explanation of the biological system's behaviour. The former status has been used in the previous chapter to include robots as *examples* of simple perceptual systems; the latter aspect will be the emphasis of much of the remainder of this thesis, and will be further discussed in section 3.4.

Conceiving of 'Autonomous Agent AI' research as modeling means that the question:

- What can this research tell us about perceptual systems?

can be linked to the issues of reasoning by constructing and manipulating models analysed above. In other words, the extent to which this research helps explain perceptual systems depends on the extent to which it satisfies the steps in that chain of reasoning:

- How accurate is the representation?
- What results does it produce?
- How can those results be interpreted?
- Can that interpretation be properly compared to the behaviour of the target system?

These questions can be related to the analysis of important aspects of perceptual systems discussed in the previous chapter:

- Do these AI models represent the specificity of behaviour; the importance of the agent-environment interaction; biologically plausible mechanisms; appropriate roles for learning and evolution?
- Can the results be justifiably compared to the operation of real (particularly biological) perceptual systems?

## 3.3 Limitations of current research

There is a wide range of levels of detail of models in this field, from those strictly based on accurate representation of biological systems, to simulations of abstract interactions the results of which are loosely described as 'life-like'. For the first extreme there are many examples of models of *parts* of sensory-motor systems, which I will not discuss here: consideration will be limited to those models that do link sensing to behaviour, in other words, complete perceptual systems, in the sense given in the introduction. At the other end of the scale, the relationship of 'life' simulations to the problems of perceptual systems is generally too remote for discussion in these terms to be profitable: the aim in such systems is often to explore the emergence of complexity from many simple interactions, which may be a significant principal for perceptual systems but does not, in itself, address the question of *what* interactions might support sensory control of behaviour.

However there is a significant amount of research that claims to address the problem of controlling behaviour in autonomous systems, to which the questions raised in 3.2.3 can be applied. Much of this research fails to adequately answer these questions. In the following sections I will discuss some of the more general flaws.

### 3.3.1 Too general

Work in this field often takes the form of 'behavioural simulation': that is, using computer programs to simulate the actions of an agent at the behavioural level, a level of description above the details of the physical mechanisms of sensory transduction, neural processing and motor implementation. The term 'animat' has been coined (Wilson, 1985) to describe the agents in these simulations, although both 'creature' and 'robot' are also commonly used.

Many examples are discussed by Meyer and Guillot (1991). Animats rarely represent any particular animal but instead have capacities like 'moving about',

'eating', 'detecting obstacles' that are considered to be generic animal abilities. An example is Coderre's 'Petworld' (1989) "a system for modeling non-species-specific behaviour" in which "Pets ... inhabited a two-dimensional, limited cartesian plain [i.e. are specified as positions on a grid] ... have a body orientation ... a limited field of view, typically 90 degrees [detects the content of one neighbouring grid position] ... can carry one rock at a time ... trees are food sources, and pets are browsers". There are a number of similar systems. A large proportion are concerned with learning or evolving agents. Others are used to investigate 'action-selection' schemes (eg. Tyrrell and Mayhew, 1991).

The validity of the concept of a 'generic animal' is hard to justify, however, when faced with the astonishing variety of species in biology. The likelihood of general mechanisms for behaviour, common to all creatures, is doubtful: there are often similarities, but also huge differences (there is little to compare between an ant and an elephant), and studying the similarities *and* differences, as in comparative biology, is important to understanding of the mechanisms. Further the possibility that these 'models' may help in explaining behaviour in real creatures is very tenuous, because these kinds of tasks have many potential solutions. In other words, even if it is shown that the hypothesised system does produce the right behaviour, this is not particularly good evidence that any real creature produces the behaviour this way. The task 'collect food' is likely to have more possible solutions than the task 'collect food from flowers with the efficiency of the honey-bee'.

In fact, the importance of task-specificity that was noted in the previous chapter for robotics and biology suggests a fundamental conceptual flaw in the attempt to deal with an 'abstraction' of the problems faced by a perceptual system. Rather than representing 'general' tasks, these simulations in fact represent *extremely simple* tasks. It is often fair to say that the tasks are so simple that their identification with biology is quite arbitrary—they more accurately represent video games. Finding efficient solutions generally exploits the task-specification; consequently the solutions found in these simulations will probably be highly dependent upon the simplicity of the tasks. This is very likely to occur when the mechanisms are hand designed; it will almost inevitably occur when the mechanisms are 'evolved'.

### 3.3.2 Overuse of adaptative mechanisms

#### Evolution

It seems obvious that the competent sensory-motor control of creatures has resulted from evolutionary adjustment, although the common conception of evolution as incremental internal change to improve fit to a fixed external environment is somewhat flawed (as discussed in the previous chapter). Nevertheless, it seems reasonable to adopt the principle of random change being passed on according to the success of operation within a set of constraints as a means of finding a good solution to the problem those constraints pose—this process has been adopted as a problem solving method in various applications, many having nothing to do with biology. On the other hand, if description of the constraints immediately suggests the optimal solution there seems little point in going through this lengthy process to obtain it.

However the possibility of using this approach for creating and understanding perceptual systems (as advocated by Koza (1992), de Garis (1992) and even recently Brooks (1992)) has some serious limitations. These are most obvious in systems such as Wilson's animat (1985), which, he claims, is concerned with "considering basic problems that simple animals must solve, and constructing behaving models aimed at solving them". His simulation involves the mutation and selection of rules governing movement about a 18x35 grid, based on input vectors that code the (three) possible states of adjacent squares. The states are called 'food', 'obstacle' and 'space' and the moves are described as the 'actions' of a 'creature' that wants to maximise the 'consumption of food'.

The basic problem is that such simulations are *not* imposing the kinds of constraints that led to the evolution of sensorimotor control systems: nor are the solutions produced by these simulations interesting or revealing when considered as examples of perceptual systems. While such work may be useful with regard to the mechanisms of evolution itself, it is not really concerned with the problems of autonomous behaviour.

Furthermore, it seems illogical to suppose that we could 'artificially evolve' creatures that satisfy the constraints of acting in the real world any faster than natural evolution did. Simulating the situation to speed up the evolution won't work, because it will only go faster if the situation is simplified, and simplified constraints will yield too simple a solution (see previous section). Futhermore, if we are currently unable to fully understand the workings of mechanisms produced by natural evolution, then a comparable system produced by artificial evolution might be as hard to understand, and thus may not contribute to explaining such systems.

### Learning

Similar difficulties apply to many learning mechanisms that have been used in implementing 'autonomous' behaviours. If the creature starts with completely random connections, the complexity of the behaviour it learns will be limited by the complexity of the task it faces, and for many current systems the tasks are too simple for the resulting devices to be of any interest as perceptual systems. In fact, the use of neural nets (despite having originally been inspired by the network of neurons in biological nervous systems) is really independent of attempts to model biology: that is, they are sometimes used in modeling biology and behavioural mechanisms, but they are equally used for many other purposes; the algorithms themselves are not particularly biological (Crick, 1989) but rather more accurately described as statistical (Kay, 1992). Geman *et al* (1992) argue from this basis that getting working perceptual systems *requires* that prior bias is built into the learning systems, and that the nature of this bias is the relevant problem to be solved—"the paradigm of near *tabula rasa* learning (i.e., essentially unbiased learning), which has been so much emphasized in the neural computing literature of the last decade, may be of relatively minor biological importance".

Not all explorations of learning in autonomous systems is quite so unstructured, but learning mechanisms seem to be used to an extent out of proportion with the role of such mechanisms in perceptual systems in nature. Even some of those whose work modeling specific biological systems has involved largely fixed

mechanisms (Cliff (1992) approvingly quotes Beer (1990)—"Only once the proper neural architectures for controlling the behavior of autonomous agents have been uncovered can we begin to examine ways in which the selective introduction of plasticity will increase the flexibility of the resulting controllers") have since gone on to suggest that using learning and evolution is a suitable means for finding these architectures (Beer and Gallagher, 1991; Cliff *et al* 1992).

There seems to be some confusion in this field over the meaning of 'adaptability'. It is seen as a desirable characteristic: "In a changing, unpredictable and more or less threatening environment, the behaviour of an animal is adaptive as long as the behaviour allows the animal to survive" Meyer and Guillot (1991). However, adaptability in this sense does not necessarily require adaptability in the sense of an ability to change internal mechanisms in response to environmental changes: an agent may be able to survive because it has fixed internal mechanisms that are not adversely affected by the changes occuring in its normal environment (no mechanism is likely to survive *every* possible environmental alteration). Nevertheless, justification for exploring adaptability in the latter sense is frequently based on the desirability of adaptability in the former sense. Thus Kaebling (1992) "One motive for making artificial agents adaptable is that many natural agents are adaptable".

This is not to say that learning (permanent change due to experience) is not a desirable feature for autonomous agents: it is simply not obvious that it is a necessary feature for basic perceptual competence. I have already argued that evidence from simple biological perceptual systems tends to bear this out.

### 3.3.3 Impoverished environmental interaction

Part of the criticism of adaptive mechanisms above is that the tasks posed are over-simplifications. Relating this to the analysis of modeling, the issue is the validity of the representation (step i*). Any computable representation of a biological system will require a large number of simplifications, and these imply a large number of assumptions about what are the relevant factors to represent, and what constitutes an adequate way to represent them. The strength of the basis

for these assumptions is critical for the strength of any conclusions that can be drawn from operations on the model. However, very few simulations in the field of autonomous agents give adequate justification for the assumptions implied by their representations; indeed the assumptions are often not explicitly mentioned. Rothenberg (1989) suggests this may be true of AI generally: "Models are generally assumed to have an analogous or imitative relationship to some real world phenomenon or system, though this assumption is often implicit. Even where explicit, this assumption usually remains vague and intuitive".

The previous chapter was intended to establish a basis for assumptions about what is critical for perceptual systems. One of these was that physical details of the interaction of sensors and actuators with the environment are critical determinants of the internal structures supporting perception. However, most current modeling involves massive simplification of the sensory and motor interaction with the environment—the examples given above are illustrations of this, and many more can be found in the proceedings of Artificial Life conferences (Langton, 1989; Meyer and Wilson, 1991; Varela and Bourgine, 1992).

What this means is that most of the work is being done on abstract problems that simply fail to represent any real problems of devising perceptual systems. Few people working in robotics would agree when Lesperance and Levesque (1990) claim about their simulation (a 'world' in which all that moves is the 'robot' which can turn 90 degrees or move forward a square, and pick up and put down 'objects') "It should be clear that in spite of the simplicity of this domain, it contains analogues to a large number of problems encountered in planning actual robot navigation, manipulation and perception". This example is typical in that the physical capacities are the least detailed part of the model. This may make the solution more complex than necessary because physics can't be exploited, or may result in a solution that is unlikely to be implementable in any real physical system, or both.

Models in this field do range from the highly abstract to rather more detailed attempts that include some representation of sensors and actuators (for example, not just 'seeing' an obstacle but getting 'sonar sensor' distance readings (Koza,

1992)). Most of the models that do include detail about physics are simulations of *actual* robots, and thus assumptions are based on experience of what is important in the real system. There is also a greater tendency to include such detail when modeling particular, rather than generic, biological examples. Obviously, it becomes easier to validate a model as an adequate representation of a system when the target system is well specified, rather than some vague concept.

Beer's 'artificial insect' (1990, Beer *et al*, 1990) is often referred to as an example of such detailed modeling, and it is true that, relative to most autonomous agent modeling, it does represent a more realistic sensorimotor situation. The agent has "the basic body plan of many insects" *i.e.* six legs, antennae and a mouth, and forward movement is determined by leg gaits driven by a neural circuit based on biological research, as is the 'appetitive' rate of mouth movements, representing feeding rates. On the other hand, turning, though described as "allowing the legs to apply lateral forces to the body" is exactly equivalent to rotating the direction headed by a certain amount, rather than a plausible way of achieving a six-legged turn, such as one side stepping faster than the other (Cruse, 1991). The turns are determined by the antennae sensing the exact angle of contact with an obstacle, or the difference in strength between the antennae of an 'odor' gradient that falls off as exactly the inverse square of distance from food: neither of these sensory situations represents a realistic situation. Only the forward motor gaits have been subsequently implemented in a robot (Beer *et al*, 1992).

While limitations in realistic representation can be justified in simulation (after all the reason for simulation *is* to test mechanisms in a simpler way than by building them), the problem with much research in autonomous systems is the failure to acknowledge the effects of these limitations when drawing conclusions from the results. If, as I have argued, the interaction with the environment determines perceptual systems, then failing to deal with this interaction in model building means that one of the most fundamental problems of autonomous systems is simply not being addressed. Few researchers, in practice, are "refusing to abstract away the problem of adaptively coping with an actual environment" (Beer, 1990).

### 3.3.4 Little result evaluation

So far, difficulties with the validity of representation (step (i*) in 3.2.2), and with reasoning from the behaviour of simulations to real behaviours (step (ii)) has been discussed. The other aspect of modeling in science identified above was the need to evaluate the model's behaviour to determine if it does constitute an analogue of the desired behaviour (step (ii*)). For models of perceptual systems this requires systematic testing of the behavioural capacities displayed by the device and realistic assessment of the comparability of those behaviours to the behaviour of real autonomous systems such as animals. It may be true that "As ALife becomes a scientific discipline, it is evaluating its organisms in terms other than simple entertainment value" (Belew, 1991), but alternative forms of evaluation are not well developed.

The problems already discussed with simplistic simulations apply to this area as well. It does not seem appropriate to describe a moving grid position that ends up in the correct state as a demonstration of a mechanism for feeding. Simple simulations are sometimes justified as attempts to discover what range of behaviour can result from the most basic of mechanisms—while this may be interesting, there should be great care not to overinterpret the results. Even though some researchers may deny any wish to make claims about biology, this is obscured by the use of terms like 'animat', 'bug' or 'organism' for simulations that have only the most trivial behavioural similarity to real animals (Wilson, 1991; Horn, 1992; Cecconi and Parisi, 1991).

On the other hand, where there is a higher degree of behavioural complexity, there is often inadequate analysis of the behaviour. Frequently 'results' consist of anecdotal accounts of one or two runs, with little indication of the characteristics, consistency, failure modes and so on that would enable an evaluation of whether the mechanism really works, under what conditions it works, and whether it is in any way comparable to the target system (Brooks' work tend to suffer from this defect). Additionally, there tends to be little comparative evaluation between alternative mechanisms for performing similar tasks.

The problems with testing are exacerbated when the target is ill-specified. That is, if it is unclear exactly what the model it is supposed to represent, then it is hard to determine what criteria the model should be tested against, or indeed what constitutes successful behaviour by it.

There can be difficulties in making comparisons with biology, because animal behaviours are rarely simple to describe, and often the description is idealized to fit a model. Reproducing this idealization needs to be distinguished from reproducing realistic behaviour. For example, Beer (1990) compares the gaits produced by his walking circuit to those "described by Wilson for natural insects" but the pictures of 'insect gaits' come from Wilson's (1966) *hypothesis* of varying overlap of metachronal waves. Thus Beer's model demonstrates the internal consistency, rather than the validity, of this theory of insect walking.

## 3.4 Methodological Proposals

The discussion above has focused on problems with the current status of modeling of biological perceptual systems in AI. These criticisms were based on the questions raised by the analysis of modeling, and the characterisation of sensorimotor mechanisms given in Chapter 2. The same considerations can be used to support a particular methodology for applying the model-building techniques of AI to the problems of understanding real perceptual systems.

### 3.4.1 Model real biological systems

Closer links to real biological systems could potentially address many of the limitations that have been discussed above: in particular, problems for representation and testing that are due to the lack of a well defined target system for the model. Biological modeling of this kind would seem to have much potential, not only as a solid basis for exploring potential mechanisms for sensorimotor control and autonomous behaviour, but also as a means of better understanding the biological systems themselves.

Modeling biological perception is not a new undertaking: within biology, building computational models is a fast growing field (see for example work reported in Koch and Segev (1989)). The extent of this modeling activity is in no way reflected by the (lack of) effect it has had on autonomous systems research in AI. It is true that a large proportion of this work is concerned with partial systems such as particular sensors. Nevertheless, it is notable that there is a far higher concern with validition and evaluation than is generally found in AI.

A good example of research with a strong biological connection that also reflects robotic concerns is the fly-vision based obstacle avoidance mechanism described in Franceschini *et al*, (1991). This biological system has been extensively studied (Franceschini *et al*, 1989) and the model copies in detail the physical receptor layout and retinal processing of the fly, using this to drive the movement of a robot to avoid collisions. It demonstrates that biological findings can be adopted for effective robotic technology.

However, it is rarely the case that "neural network architectures abstracted from biological systems can be directly applied to the control of autonomous agents" (Beer *et al*, 1992). In fact there are very few biological systems, especially complete perceptual systems, that have been studied in anything like the detail required for translation into electronics and programming code. Looking at biological systems by attempting to build models of them is just as much a potential opportunity for AI to contribute to biological understanding of perceptual systems.

Jamon (1991) provides interesting examples of stochastic models that can account for animal behaviours usually attributed to mental maps, and notes

> "Simulation models can therefore considerably change our views about animal orientation. They allow performances to be attributed to alternative hypotheses. In view of the strong anthropomorphic tendency pervading the study of animal orientation, they can serve as a heuristic tool for reconsidering long distance orientation theory."

In other words, the requirements imposed by model building may lead to increased appreciation of the non-representational nature of perception:

> "AI must ... incorporate what can be learned from a study of nervous systems. It must stop reasoning by analogy with well studied, but irrelevant, physical systems ... and must instead reason by analysis of relevant facts about biological systems that actually have intelligence. This change will require AI to abandon the notion of intelligence as a purely abstract information processing activity" (Reeke and Edelman, 1988)

### 3.4.2 Use physical models

Given a specific biological system to model, it becomes more viable to validate models because there is a basis for assumptions about what should be represented and how. A number of arguments can be made for using robots as substantial analogues of perceptual systems, rather than computer simulations as formal analogues (many are raised by Brooks (1986)). This is not to say that computer simulation is *necessarily* a bad method of modelling: rather, there are reasons why achieving a representation can be better done by using a physical model for the sorts of problems perceptual systems face.

Cliff's (1990,1992) 'SYCO' simulation is a useful example to illustrate some of these points. The target system is the visual tracking mechanism of the hoverfly. The visual signal is modeled in some detail, using optical ray tracing and a retinal receptor distribution model to determine the input to the neural layers. The motor response is more simply represented as turning, without details of how flight control is established, but the result of this movement is then used directly to determine the visual input, modeling "the external feedback loop provided by the environment". Such detail is computationally expensive and requires a great deal of parameter tuning: in the experiments given in (1992) the ray-tracing is replaced by projection equations for tractability. On the other hand, validity

as a biological model is in several places *disclaimed*, allowing some biologically implausible elements to be used.

If biological accuracy can be sacrificied to some extent, then there seems no reason not to use *real* sensors that, embedded in the real world, by-pass the need to *calculate* the effects of the world. There seems little to distinguish between the complexities of ray-tracing and the "problematic engineering issues of sensorimotor transduction" that Cliff claims to avoid by using simulation. It is often easier to model physical systems with other physical systems than with mathematical equations: transformations corresponding to movement will contain the same information and the same sources of noise and error; moreover this will occur in parallel, and not occupy processor time.

Furthermore, using a robot helps in validating the system. If the model works in real world interactions then it is not necessary to show that the results are *interpretable* as sensory-motor behaviour, but only that it is *comparable* to the target set. The latter is still not a trivial task, but at least there is one less level of justification required.

An objection that Cliff raises to using *real* real world systems for modeling is the problem of repeatability—"in simulated 'real' worlds *identical* conditions can be recreated as many times as required". This problem does exist, but it is worth recognising that it is also faced by those who study real systems in biology and psychology—and these fields have been developing methodology for studying behaving systems for at least a century, which can be adopted for analyzing robot behaviour. Moreover, repeatability—behaving the same way under identical conditions—is generally a less interesting characteristic in perceptual systems than robustness—performing adequately under a wide variety of conditions. Robots are more useful for testing the latter.

### 3.4.3 Establish evaluation methods

The previous point obviously ties into the issue of evaluation. Once again, having some specific biological system as a target is advantageous, because comparable

methods can be adopted to measure the behaviour, and direct comparisons drawn. This is useful both for evaluating the effectiveness of the mechanism that has been implemented; and for generating hypotheses and predictions for the biological system being modeled.

An illustration of this is the work of Arbib and fellow researchers (1987,1989) on a neural model of the toad's visual-motor coordination in moving towards and snapping at flies and worms. The model is based on both behavioural and physiological studies of this system (Ewert, 1987), and aims to devise "processing schemes that could plausibly be carried out in neural structures"(Arbib and House, 1987) known to exist in the frog's visual system. Although this work does simplify the animal-environment interface, the simplifications are at least partially rooted in observations from the real system, and the results are closely compared to biological results. Consequently the system is far more convincing as a useful investigation of perceptual control than most computational simulations.

Experimental design and statistical analysis are not often well employed in AI (Cohen, 1991). Yet such techniques are necessary to validate systems; particularly systems with complex behaviours that can't simply be classified as 'working' or 'not working'. There needs to be decisions about what is relevant to test, plans for how it can be tested, and criteria for analyzing the results; all *before* the system is deployed. It is very easy to pay attention only to the best runs of a robot if the methods of evaluation have not been considered in advance.

More generally, the points made above about the nature of modeling must be borne in mind: the conclusions drawn from a model are limited by the accuracy of the model. Appreciating the strengths and weaknesses of a particular representation is an important aspect of evaluation.

> "If due caution is utilized, the employment of model systems can be a superlatively powerful weapon for both qualitative and quantitative investigation. But without proper understanding of the essential nature of modeling relations, it can lead us far astray." (Rosen, 1983)

## 3.5 Conclusions

It often seems, in this field of research, that while the above ideas (modeling real biology, using physical models, applying experimental procedures) are accepted as being *desirable*, they are not generally considered *viable* for current projects. The rest of this thesis describes the implementation and evaluation of a robot model of cricket phonotaxis. Thus by demonstration, I hope to establish that the approach suggested, though not easy, is viable; and moreover produces results that make the difficulties (of studying a real biological system, of constructing a robot, and of testing and analyzing its behaviour) worth overcoming.

# Chapter 4

# Cricket Phonotaxis

## 4.1 An insect sensory-motor system

The particular biological example of a perceptual mechanism that is modelled in this thesis is phonotaxis in the cricket: the ability of the cricket to find a mate by moving towards the sound it produces. The behaviour and its neural underpinnings are one of the insect systems most thoroughly studied in neuroethology, and thus there is a good deal of experimental data to draw upon; yet the fundamental mechanism by which the cricket uses sound sensing to control its behaviour is still not well understood, which means that modeling may make some contribution.

The first section of this chapter contains some general comments about sound location and the task of phonotaxis for biological creatures. The remainder of the chapter will deal with this behaviour in the cricket (*Ensifera: Gryllidae*) only. More details of the behaviour, the anatomy and properties of the auditory system, and of experimental methods will be provided as background. Then some hypotheses about the mechanism will be discussed, focused by the application of the principles expounded in Chapter 2. A specific mechanism is proposed as a simple yet viable explanation of taxis, that can be further investigated in a robotic model.

## 4.2 Finding sound sources

### 4.2.1 Directional properties of sound

Animals can find sound sources because of the way that sound is propagated in the environment. In the case of an ideal source pulsating at a fixed frequency, molecules of surrounding fluid are cyclically compressed, resulting in a sinusoidal changes in pressure: thus waves of compression are spread in the surrounding medium. From a different position in the medium, the location of the source can be determined by detecting properties of the passing wave. Near the sound source, the particle movements (which are inherently directional) contain sufficient energy to be detected. In the far field, the amplitude of the wave is proportional to distance from the source (for the ideal case, it decreases as $1/distance^2$, but in most natural cases environmental factors will alter the attenuation rate), so differences in amplitude can indicate direction. The phase of the wave also changes with position in the sound field (cyclically rather than monotonically), and so again, differences can be used to derive the direction of the source. The spreading of the wave can be modified (for example blocked by a solid object) and so the direction of the sound relative to the receiver/modifier orientation can be determined.

The latter method may sound complex, but in fact is commonly found in animals. The human head distorts the sound field so that the ear on the side near the source receives a louder sound than the opposite ear. Many pressure sensitive receivers are shaped so that that identical waves from different directions will affect them differently; for example, a diaphragm (unlike the 'ideal' point reciever) will respond differently to waves that arrive in a direction normal to its surface and waves that arrive at an oblique angle to it. This means of determining sound direction contributes to many animal systems that use comparison of amplitude as the basic method of determining sound direction. With two ears, amplitude difference specifies a region of possible directions: three measurements are needed to specify direction uniquely. Most animal systems obtain the additional measure-

ments by movement (*i.e.* over time) rather than developing a third receiver (*i.e.* over space).

Phase information can be used in a similar manner to amplitude, that is, by measuring the difference between receivers, although the same directional ambiguity with two receivers occurs. The time difference between ears is only in the order of microseconds, and detecting it neurally requires dedicated processing arrays, such as found in the owl (Konishi, 1993). The process is further complicated if the sound source emits multiple or varying waveforms.

Particle movement detection is often found in insects and other arthropods (Ewing, 1989). For example, the cerci of the cockroach are made up of hairs that articulate in different planes, thus large deflection of particular hairs indicates that the sound wave is moving in a particular direction. As the effective range of particle movement is further in water than in air, it is not surprising that similar mechanisms have been found in fish.

### 4.2.2 Orientation and Taxes

Given sensors that enable detection of directional properties of sound, an animal can make orienting responses to sound, ranging from moving towards or away from the sound, through turning the body or a body part to face it, to (in humans) saying 'it's over there'. Saying an animal can 'locate' a sound generally implies nothing about its subsequent behavioural response to it: it may be treated as the animal extracting a piece of information (the direction of a sound) from its environment; but as stressed above, describing the *complete* perceptual system must include the subsequent response.

Taxes are movements in response to sensory gradients. By following a sensory gradient, animals can come into contact with the emittor, and thus can encounter things important to their survival, such as food, a dark place to hide, or a mate, more efficiently than by undirected wandering in their environment. Taxes are found in many animals, from amoebae upwards. They are a particularly robust way of guiding movement: continuous reactive course correction is simpler to

implement, and less prone to error, than accurately identifying the direction of the desired source and planning a path towards it. Braitenburg (1984) describes a range of interesting behaviours that can be built up by combining simple taxis traits. On the other hand, they require a more or less continuous signal from the source, propagated in the environment in an appropriate manner, and thus are only suitable for finding certain kinds of things.

Phonotaxis for finding a mate is found in a number of species including other arthropods (grasshoppers, cicadas), amphibians and birds. It seems a suitable capability to study as a simple perceptual system, as it has many of the features noted in Chapter 2: it is usually quite specific to particular sounds; central representation seems unnecessary; the physical properties of the receptors and actuators are significant; movement in the sound field disambiguates sensor information; it is genetically fixed rather than acquired behaviour; and, though simple enough for the whole system to be grasped, sufficiently complex that the actual mechanisms in particular animals have not yet been determined.

## 4.3 Background

The body of research on cricket phonotaxis is extremely large and varied; and there is no adequate recent overview. General papers that cover important behavioural and physiological results include Huber (1983), Huber and Thorson (1985), and Schildberger (1988); more background about the ecology is given by Ewing (1989).

### 4.3.1 Ethology

The female cricket can find a conspecific male by walking or flying towards the 'calling song' the male produces. This sensory cue is sufficient, but not necessary, for finding the male: there is some evidence that in the natural situation, the female will use other cues if they are available, such as setting a directed course using visual cues (Stout *et al*, 1987), or if the incidence of male crickets is sufficiently

high, random search (Hissman, 1990). Using only auditory cues, however, the female may cover quite a large distance—ten to twenty or more metres—usually negotiating uneven vegetation-covered terrain, yet can reliably find one calling male, despite other males (and other sounds) in the vicinity. Once in contact, further sound cues (the 'courtship song') and chemical cues determine whether the pair will mate.

The calling song produced by the male cricket is species specific, although not all songs of all species are distinct. The crickets forewings have a rasp and file: by drawing one wing across the other, wing surfaces are made to vibrate at a regular frequency. Having resonant surfaces, the resulting sound is an almost pure tone (typically in the range 4-5kHz). Each down-stroke of the wings will produce a short burst (10-50ms) of sound, followed by a silent gap during the upstroke of roughly equal length. These bursts (known as syllables) may occur in groups (known as chirps) separated by longer pauses. It is this chirp structure, at rates of less than 10Hz, that is the pattern in the sound audible to humans. Figure 4–1 illustrates these song features for two species often used experimentally.

There is evidence (Simmons, 1988) of males aggregating in one area to sing together, thus increasing the chance of attracting females to the area, but maintaining a certain distance (around 1 metre) from each other, possibly by monitoring the relative intensity of nearby songs. The singing is prevalent at certain times of the day (different singing times may also separate species that share an area (French and Cade, 1987)), mostly after dark (to minimise predation) and may last several hours.

Females only perform taxis in the adult stage of development (crickets develop through a series of nymphal stages rather than having a larval form like most other insects). Encounters with males ultimately reduce responsiveness to the song, probably through changes in hormone levels.

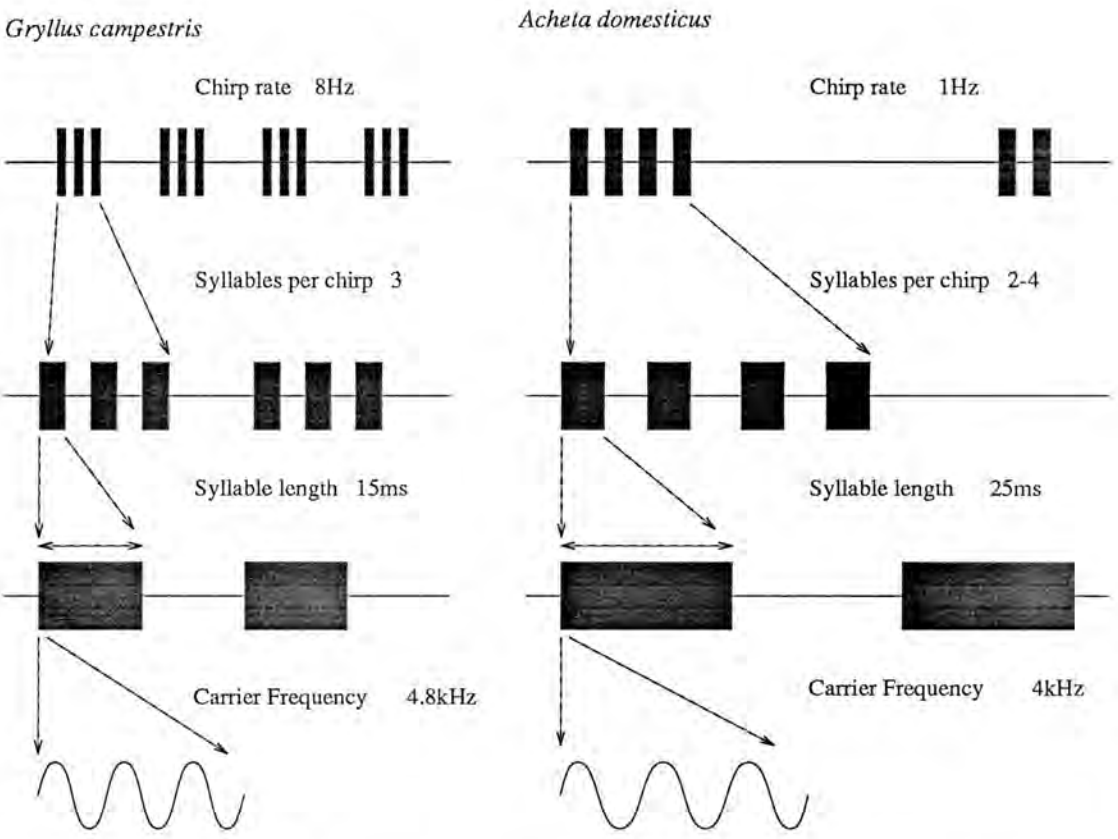



**Figure 4–1:** Hierarchy of song features for two cricket species

### 4.3.2 Experimental methods

It is understandably difficult to track an insect through vegetation at night, so there are few field studies of cricket behaviour that provide more than the general details mentioned above. Instead most behavioural results have been obtained under laboratory conditions. The results to be discussed below require an understanding of the most common methodologies employed.

#### Moving

The earliest demonstration that sound was a sufficient location cue for the female cricket used the newly invented telephone to send the male calling song to a female, which approached the earpiece (Regen, 1913). Using a loudspeaker to replay a recorded male song or a synthesized song is still the stimulus employed in most experiments. Many of the earlier experiments were conducted by placing a speaker

at one arm of a Y-maze, and using the frequency of choice of that arm to indicate phonotactic preference.

An alternative is for the cricket to be released into a walled arena and allowed to run freely towards the sound source, or between two sources. This allows various characteristics of the path taken towards the sound to be more closely examined, as the track is not constrained as it is with a maze. Though this does by and large eliminate the natural complexity of the cricket environment and the sort of sound distortion that occurs in it (the arena floor is empty, and it is surrounded by sound absorbant material), this method does at least maintain the normal relationship between perception and movement for the animal, which the following methods do not.

### Tethered

One perceived disadvantage of the arena paradigm was that it was very difficult to control the exact sound-field of the cricket, as this was dependent on how it moved: furthermore, the cricket constantly had to be captured and rereleased to build up data. It was difficult to compare these behavioural results with the results of neural recordings, which could only be made on a motionless animal.

An alternative method for examining the behaviour was to fix the cricket (usually waxed to a wire) above a ball or frame that it could turn with its feet when attempting to walk. Thus the cricket's position remained fixed, but the movement of the ball could be used as a continuous record of the direction in which the cricket was attempting to move when presented with certain sounds at certain angles. An advantage is that the walking behaviour can be measured over long periods without being terminated by the animal reaching the source: a disadvantage is that the animal is not getting the usually environmental feedback for its movement, that is, turning movements of the legs don't result in the normal movement of the auditory organs in the sound field. This prevents investigation of how the animal may vary its direction when approaching a sound.

There are some other techniques used that are based on essentially the same principle. Stabel *et al* (1989) use two wheels, one for the legs on each side, and take relative speed as a measure of turning tendency. For crickets that usually fly, the bending of the abdomen of a fixed animal can be used to indicate attempts to turn in flight (Pollack, 1986).

Schildberger and Horner (1988) have succeeded in getting behavioural data from a tethered cricket while simultaneously recording from its nervous system: one of the few cases in neuroethology where a direct link, rather than an inference, has connected measurements of neural functioning to particular behaviours.

### Compensated

The most extensive studies of cricket behaviour have used the treadmill paradigm developed by Kramer (1976). In this, the cricket walks freely on top of a large spheroid treadmill. A tracking camera above the sphere causes it to move in automatic compensation to keep the cricket on top of the sphere: the direction and velocity of the cricket's walking can be recorded from the compensation required. The main advantage over the previous method is that the animal can change its direction relative to the sound, which makes it more closely resemble the normal situation: however, it still doesn't actually approach the sound. Crickets may walk continuously on the treadmill for an hour or more. Some limitations with this methodology are discussed by Weber*et al* (1981).

### General comments

It should be noted that despite the development of these techniques, it is still quite difficult to carry out behavioural experiments, as there are a large number of conditions (such as light levels, acoustic surroundings, chemical cues, temperature, source of animals, handling of animals etc.) that can affect experimental success. At an early stage in this thesis I expended some effort to establish equipment and procedures for working with actual crickets: but only one animal showed any signs of phonotactic response to a recorded song, with all others confining their

behaviour to attempts to escape the arena. Most reports on behavioural experiments come from just a few particular laboratories, which have well established experimental programs for cricket research.

It should also be noted that there is a great deal of individual variability between crickets: for example, Weber *et al* (1981) report that of 31 crickets caught in the wild, only 7 tracked well on a treadmill, 7 did not track at all, and the rest ranged between these extremes. Most experiments start by identifying 'good trackers' and use only these animals in the subsequent tests; a 'strong' result might be a behaviour exhibited by just over half of these. Many of the most interesting experiments lack confirmation, and apparently contradicatory results are not uncommon.

### 4.3.3 Neurophysiology

Neurophysiological experiments are generally more difficult (and more dependent on sophisticated techniques) than behavioural ones, and many of the above comments apply to the neurophysiological details discussed below. Only the peripheral auditory system has been examined in any detail; results for more central neural processes are very sparse.

#### The auditory organs

The auditory organs of the cricket are located in each foreleg tibia. Each leg has two tympani, a small (1mm) anterior tympanum and a larger (6mm) posterior typanum. Their vibration is caused by the sound pressure reaching them, which comes both directly from external sound waves, and indirectly through an air-filled tracheal system that connects the two ears to each other and to two further openings on the sides of the thorax (the spiracles) (Figure 4–2).

The interaction of the sound waves arriving by different routes causes phase cancellation, which depends upon the the angle of incidence of sound. Consequently the vibration of each tympanum is proportional to the direction of the

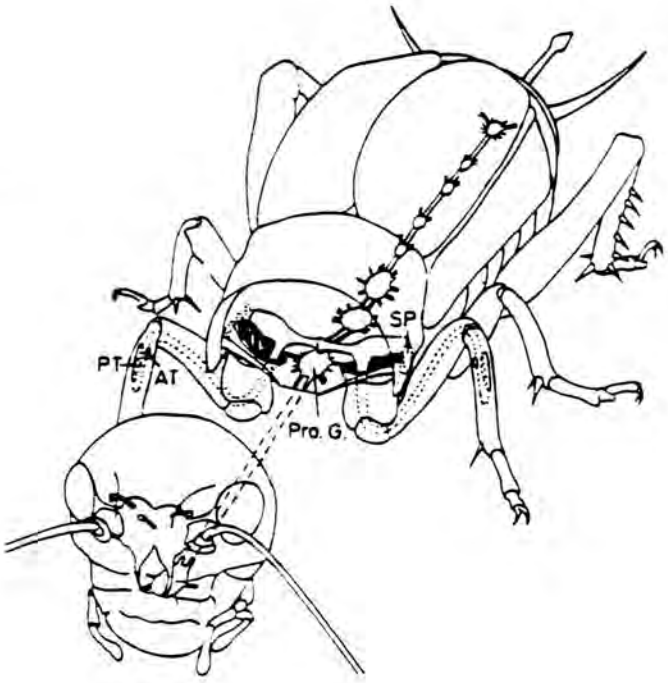



**Figure 4–2:** The auditory organs and tracheal system of the cricket: showing posterior tympanum (PT), anterior tympanum (AT), spiracle (SP) and prothoracic ganglion (Pro. G) which is the first synapse site for the auditory receptor neurons. From Schildberger (1988).

sound source. The need for such a mechanism is clear when it is noted that the distance between the forelegs of a cricket (around one cm) is insufficient for detectable intensity differences, and the phase difference is orders of magnitude below the speed of the neural pulses. In 1977, Hill and Boyan demonstrated that directionality in the response "does not depend on a diminished external sound pressure at the ear ... but rather on sound transmission along the leg trachea" and they outlined a simple version of this system (see Figure 4–3). The tympani will vibrate according to the differential pressure arriving externally and internally: considering only the two posterior tympani, separated by a trachea of roughly 1/4 the wavelength of the calling song: if the sound is on the left (fig. 4.3A) the sound wave travelling through the trachea from left to right ear travels the same distance as the external wave, so they arrive in phase, on opposite sides of the tympanum, and cancel out: whereas with the sound on the right (fig.4.3B) the wave travelling back through the trachea from the left will have travelled twice the length of the

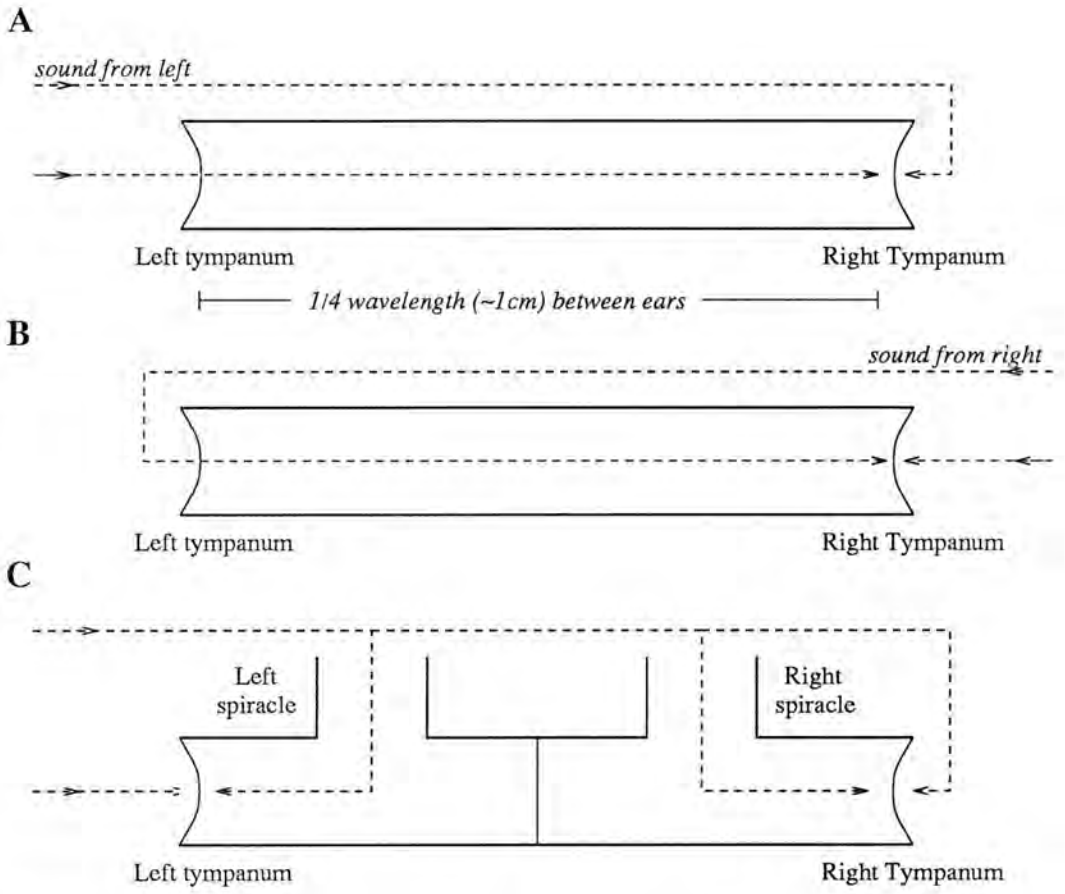



**Figure 4–3:** The phase cancellation mechanism A: sound from left B: sound from right C: spiracle contribution

trachea (half the wavelength) relative to the signal arriving at the outside, so the waves will be in opposite phase, on opposite sides of the tympanum, and hence will compound. Sounds at various angles between these extremes will cause differences in relative phase, thus different levels of cancellation, and hence the effective intensity at the tympani reflects the direction of the sound source.

In fact, Boyd and Lewis (1983) found that directionality of response was not affected by covering (with wax) the opposite tympanum, but strongly affected by covering either or both spiracles. Similarly, Larsen and Michelson (1978), measuring tympanal vibration, showed better transmission of the signal to the tympanum from the ipsilateral spiracle (100%) than from contralateral spiracle (35%) or from contralateral tympanum (15%). Thorson *et al* (1982) describe a phase cancelling mechanism that uses only the contribution of the ipsilateral

spiracle (fig. 4.3C) which is more biologically plausible than that described above. As the distance from tympani to spiracle is smaller than 1/4 wavelength, the waves will never be in completely opposite phase so the directionality effects are somewhat reduced, but still sufficient to form the basis of localization.

To be exact, the evidence seems to suggest a rather complex interaction of the sound waves, with proportional contributions from the spiracles, the opposite tympanum and also from the small anterior tympanum (Schmitz, 1985). The attenuation of the signal in the trachea depends upon the wavelength (Kleindienst *et al* (1981) suggest it is least for the calling song frequency) and the speed of sound transmission in the trachea is slower than in air. The shape of the typanum itself also contributes a degree of directionality to its response (Erulkar, 1972). So although the mechanism is understood in principle, it is as yet impossible to devise an accurate quantitative model—even estimating the length of the trachea is difficult. As Michelson (1983) notes in this context "it is often very difficult to find equations describing the acoustics of biological systems ... large theoretical models have been published which—although fitting some experimental data—have very little in common with the actual system". Fletcher and Thwaites (1979) designed an electrical analog model of the cricket's acoustical system, intended to permit calculation of "specific numerical results for frequency response, directional discrimination and other related quantities, provided that the necessary physical parameters of the auditory system are available"; but despite including many careful estimates, their results bore little relation to experimental ones.

Thus, given a description of a particular external sound field, the vibrational response at the tympanum that initiates the neural response cannot be predicted with any accuracy. For complex fields, for example, with two sound sources, even a qualitative estimate is difficult. An alternative means of determining the response is to measure it directly, recording tympanal movement itself, or the firing rate in the auditory nerve. Boyd and Lewis (1983) constructed an intensity-response curve (using a single fixed sound source of varying intensity) for the firing rate of the auditory nerve, then used this response rate to estimate the effective intensity at the typanum for sound from varying locations. They found that with

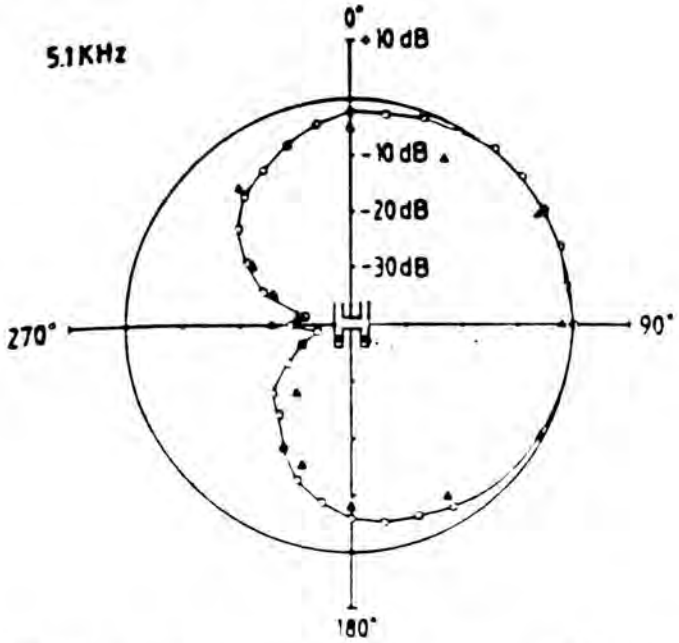



**Figure 4–4:** Intensity response to sound from different directions, measured by vibration of the tympanum (triangles) and by firing rate in the auditory nerve (circles). From Larsen *et al* (1983).

a signal between 4.4 and 5.3 kHz,the response was cardioid (see Figure 4–4). with effective intensity differences of 18-35db. Larsen *et al* (1983) report similar results from measuring the vibration of the tympanal membrane. Measurements for more complex sounds have not been published, which can make interpretation of results from complex behavioural experiments rather difficult.

### Auditory receptors

There are 50-70 receptor cells located in the trachea near the posterior tympanum that transform the vibration of the tympanum into nerve impulses (Kleindienst *et al*, 1981). The receptors vary in preferred frequency: the mechanism underlying this frequency selectivity is not known, although it appears to be a mechanical-electrical property of the receptor cells themselves, rather than a property of their arrangement as a group (Oldfield *et al*, 1986). A significant proportion are tuned to the 4-5kHz frequencies of the normal calling song, with response bands ranging from narrow to very wide. Individual receptors have a limited intensity range (from threshold to maximum firing rate) of around 20dB, but different cells have

different threshold levels so that in combination the response covers a greater range (up to 50dB)—again, the mechanism determining these different thresholds is not known.

Within its given range, the receptor encodes the intensity level of the vibration in two ways, i) the time lapse till firing starts (latency) and ii) the rate of firing. Esch *et al* (1980) found the firing rate and inverse latency to be roughly linearly related, both logarithmic to intensity up to saturation. The number of spikes per 20 second syllable ranges from 0-5+, the latencies from 10-4ms. There is no spontaneous discharge. The recovery time of these cells is around 1ms, so the normal calling song patterns (with 20-50ms syllables) are not significantly distorted by it.

The axons of these receptor cells form the auditory nerve which runs from the foreleg tibia to the prothoracic ganglion where they form synapses with interneurons (see below). Their arrangement is apparently ordered by the frequency and intensity characteristics of the receptors (Romer, 1985). It should be noted that the functioning of the peripheral auditory system as a whole takes a phase difference between the ears, converts it mechanically to an intensity difference in the tympani, and then converts it neurally to a phase difference (of milliseconds rather than microseconds) in the auditory nerve.

### Auditory interneurons

Six pairs of auditory interneurons (neurons that respond to auditory stimulation) have been identified in the prothoracic ganglion of *Gryllus campestris* (four pairs are shown in Figure 4–5).

There are two pairs of omega neurons (ON1 and ON2): so called because of the shape of their axon, which crosses the midline to connect to its pair. ON1 receives excitation from the ipsilateral auditory nerve and is inhibited by the contralateral ON1. This reciprocal inhibition enhances the difference in response between the two sides, Michelson and Larsen (1985) estimate that it can increase a firing rate difference by 60%. The latency difference is also enhanced, as again it is linearly

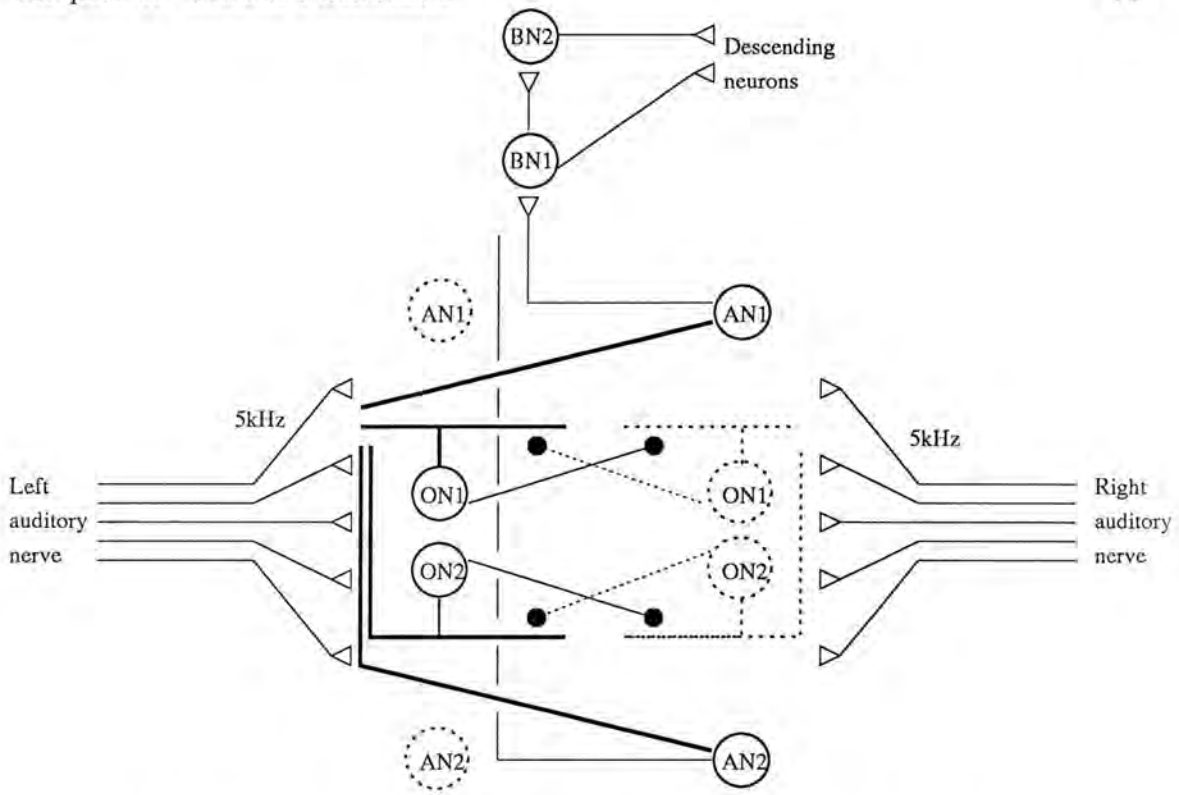



**Figure 4–5:** Schematic diagram of auditory neuron connections. Second of pair is shown dotted, with connections not shown for second AN neurons to simplify. Thin lines are axons and thick lines are dendrites. Note: AN1 and ON1 have input only from 5kHz section of auditory nerve; ON neuron pairs have mutually inhibitory links; AN neurons connect to brain neurons (BN).

related to the firing rate. Although the response properties of the neuron allow it to code most temporal patterns, Wiese and Eilts-Grimm (1985) suggest that time constants for feed-back between the cells cause the best enhancement of intensity difference to be at the syllable rate of the conspecific calling song. ON2 appears to have similar shape and similar connections to ON1 but its response properties are less clearly related to features of the calling song, and it has not been so well studied.

There are two pairs of ascending neurons (AN1 and AN2), in which the cell body is contralateral to the dendritic field, which receives excitatory input from the auditory nerve. AN1 shows preferential tuning to the calling song frequency, as it apparently connects principally with those auditory receptors tuned to that frequency. It is considered critical to phonotaxis: hyperpolarizing this neuron sub-

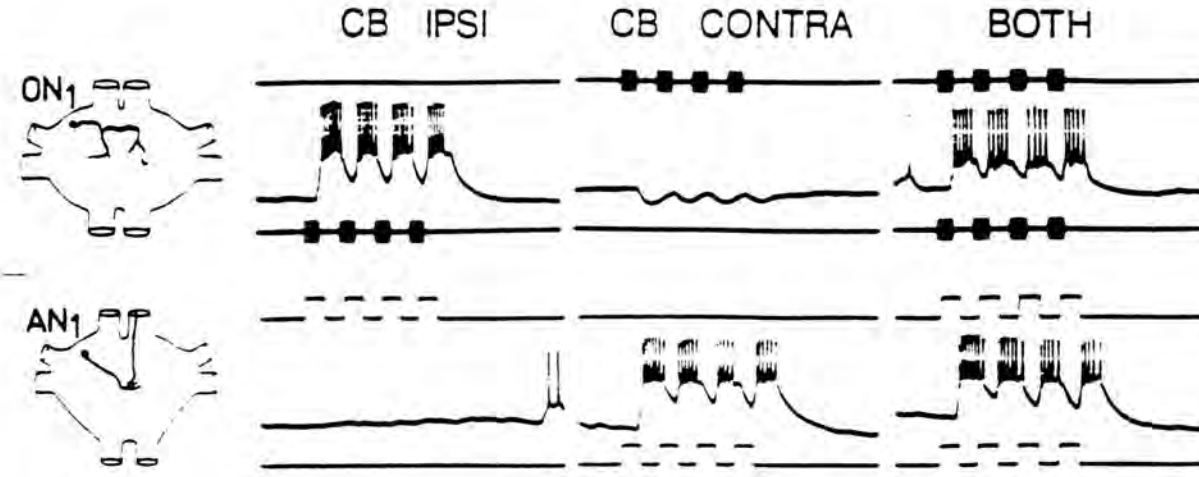



**Figure 4–6:** Traces of neural response to sound signals. Upper is ON1: excited by ipsilateral and inhibited by contralateral sound. Lower is AN1: excited by contralateral sound. Both have raised potential with superimposed spikes.

stantially alters the walking direction (Schildberger and Horner, 1988). It has no background activity. Excitatory input causes a rise in postsynaptic potential with superimposed spikes (see Figure 4–6), firing is maintained with little adaptation as long as the signal lasts. The rate of decay of potential after cessation of excitation determines the 'critical pauselength' between inputs that can be detected in the neural response: decay to 50% potential takes around 15ms. Thus the firing can reflect the pattern of the normal calling song (Wohlers and Huber, 1982). Latency from 38-16ms, and firing rate from 0-10 spikes per syllable code the intensity of the signal.

AN2 has fewer characteristics specific to the calling song: it responds to a wider range of frequencies and does not encode the pattern well. It has been suggested that it is involved in other auditory behaviours rather than phonotaxis (though it may well still have some effect on the latter)—in particular its sensitivity to higher frequencies may implicate it in response to the 'courtship' song.

Hennig (1988) notes that in all studied cricket species there are ascending interneurons that more or less fit the AN1/AN2 classification, in particular, there is always a pair (like AN1) with best frequency response at the carrier frequency

of the species calling song. For example, Stout *et al* (1988) report for *Acheta domestica* that the cell L1 has maximal sensitivity at calling song frequency (4-5kHz) but a limited dynamic range, saturating 10-15db above threshold. Below the threshold the temporal features of the calling song are reproduced, above the threshold, there is an initial spike, with latency decreasing with intensity increase, followed by a burst of firing whose rate, but not duration, increases with syllable duration.

The other pairs of interneurons (not shown in Figure 4–5) that have been identified are a T-shaped neuron (which ascends and descends) and a descending neuron. The latter appears to have a good response to lower frequencies and may be involved in escape reactions. Neither has been well studied.

### Auditory brain neurons

Schildberger (1984) identifies two brain neurons (BNC1 and BNC2) which respond to auditory stimuli and appear to be connected to the ascending neurons AN1 and AN2 (BNC2 via BNC1). These neurons respond to the calling song, but with fewer action potentials and less intensity dependence. Their response to differing syllable rates in calling songs was analyzed: they show, variously, high, low and bandpass characteristics relative to the normal rate; that is, some neurons fire only when the syllable rate is within the range to which that species will perform phonotaxis (see discussion below).

Further connection from these brain neurons to descending neurons has been noted anatomically but the response properties have not been explored.

### Motor control

In 1982, Wohlers and Huber pointed out that "the process whereby afferent input ultimately triggers efferent signals used in controlling phonotaxis and steering is not known"; this appears to still be the case at present. What evidence there is about neural mechanisms of walking in insects is concerned largely with the rhythmical maintainance of stable gaits (Cruse, 1991): starting, stopping and

turning are not well explored. It has been observed that insects can turn either by altering the stepping size or rate on one side of their body (i.e. while still maintaining forward momentum) or by reversing the steps on one side (so turning is roughly on the spot). But there are few clues as to what kind of neural signals might be required for a cricket to adjust its direction in response to sound.

Bailey and Thompson (1977) describe the path taken by the cricket in an arena as consisting of discrete moves; the animal pauses (for about 400ms), makes a turn (80% of the time towards the target) then moves in a straight line for several seconds. The stopping rhythm is apparently unrelated to the song rhythms. The animal usually stops with the ipsilateral tibia forward. In other arena studies this sort of zig-zag pattern seems less evident, although it had been noted in an early field study.

Weber *et al* (1981) describe phonotactic walking on the Kramer treadmill as 'spiky' (referring to the velocity plot): it consists of quick tripod gait walking (8cm/s) broken by brief stops (less than 500ms). Direction change generally occurs after stops, 25% of the time turning the wrong way: the animal oscillates $\pm$ 30 to 60 degrees around the speaker direction, the deviation decreasing with a more intense signal.

Schmitz *et al* (1982) in a more complete analysis of walking on the treadmill (including comparisons with and without sound) also note this oscillation, and its decrease with intensity, but dispute the dependency of turns on stopping. Changes in direction also occur while the cricket is moving forward (including turns to track a new speaker when the sound is switched). Turns while walking tend to be smaller, but this may be a motor limitation: turns after stopping may overcompensate for error angles.

These observations tend to suggest that the cricket makes a single turn to correct course direction when too large a difference between the ears occurs; that this turn may depend on input over several syllables rather than being immediate; that it then moves forward again rather than recorrecting a turn that is in the wrong direction. But the information is not conclusive on any of these points and there is really little indication of what sort of neural signalling drives these turns.

## 4.4 Mechanism

The question addressed by phonotaxis research is 'What is the neural structure that generates the observed phonotactic behaviour of the cricket?'. It is fair to say that, so far, this mechanism is not known. Taxis itself is generally explained by the principle of 'turn to the more strongly stimulated side'; the species specificity is attributed to a separate process of filtering the frequency and temporal pattern to identify the signal and thus decide whether to 'turn on' the taxis process. But as Schildberger and Horner (1988) write "It is still entirely unclear whether or how these two aspects of the calling song are processed independently of one another in the brain, or how the brain triggers and controls phonotactic walking".

There is thus an unbridged gap between the general principles proposed and the specific neural functioning that is investigated. Much of the research is purely observational (true to the atheoretic roots of ethology); what hypotheses there are tend to use terms, like 'identify' and 'decide' above, that are suggestive of cognitive processes but do little more than *describe* what is observed, rather than explain it. This lack of hypotheses makes it quite difficult at times to interpret the experimental evidence as there is no basis for organising the results.

In previous chapters I have suggested several principles to apply in understanding sensory systems: the need to appreciate the behaviour they support; the direct interactive coupling that characterises behaviour at this level; the importance of the physical nature of the transducers; and the utilization of special assumptions that are inherent to the particular situation so created. In the following sections, these ideas will be applied to the problem of phonotaxis to generate some hypotheses about the mechanism of the cricket system, and evidence from cricket research will be used to evaluate the viability of these hypotheses. They will provide the basis for the mechanism described in subsequent chapters that enables a robot to perform phonotaxis.

### 4.4.1 Inherent recognition in taxis

The behaviour of phonotaxis (moving towards a specific sound source) suggests that an appropriate mechanism would be one which can only successfully perform taxis to the right kind of signal. The cricket does not require an independent ability to locate any signal (although there may be some other auditory signals that also generate directed movement (Pollack and Plourde, 1982)), nor is it necessary for it to recognise a signal that it cannot locate. Rather, non-conspecific signals can be inherently rejected by the failure of the taxis mechanism to operate with them.

This idea arose initially from consideration of the peripheral auditory mechanism: the tracheal system that uses phase cancellation to produce the difference in intensity at each tympanum. This mechanism can only work if the signal is at a particular frequency, because the phase shift depends on the (fixed) length of the trachea. In other words, the phase cancellation won't produce the appropriate orientation-specific difference at frequencies other than that of the calling song. Thorson *et al* (1982) find that changing the frequency can result in anomalous phonotaxis, that is, tracking at the wrong angle: this does require a higher signal intensity (to overcome the neural tuning effects) and seems to indicate that rather than rejecting a signal with the wrong frequency, the animal is simply less able to find it.

Experimental separation of the processes of recognition and taxis is difficult, as the only indication that a cricket recognizes a particular sound is that it is able to walk towards it; taxis can only be examined by using recognizable signals. It has been suggested that characteristic changes in the walking mode (the 'spiky' walking mentioned above) observed in the response to signals that the cricket cannot orient towards, for example, sound being broadcast from exactly above the animal, are an indication of a separate recognition mechanism (Weber and Thorson, 1988). This assumes that the orienting mechanism does not operate at all unless a difference between the inputs exists, which need not be the case. Indeed, there is evidence that the inherent variability of neural response can result in uneven neural firing in response to a symmetric auditory input, which may well

be sufficient to cause stops and direction changes characteristic of phonotaxis, while failing to generate consistent maintainance of a particular direction.

Some studies have manipulated the auditory input so that a recognisable signal is present but the response amplitude (defined as neural firing rate) is inconsistent with the direction of the calling song. Pollack (1986) found that presenting different songs to each ear resulted in equal levels of activation in both ON1 neurons, but turning (abdomen flexion) was to the conspecific song. Stabel *et al* (1989), using a paired treadmill, presented a song from above coupled with a continuous tone horizontally from one side. This resulted in higher mean activity in neurons on that side, but taxis *away* from that side—the higher activity having obscured the song pattern. They conclude that the signal is first processed on each side for recognition, and the 'strength' of recognition on each side is the basis for comparison for location, but the results are also consistent with a taxis mechanism that depends on 'recognisable' features of the signal (see below). Doherty (1991) used a similar paradigm to Stabel *et al* (1989) but with a 'not normally tracked' calling song, rather than a tone, from the side, resulting in phonotaxis towards that side, contradicting the previous result. But the rather complex experimental design does not seem to provide clear support for his alternative assertion that recognition and location are essentially independent.

### 4.4.2 Latency comparison

How might the phonotaxis system operate so that it works only for the right kind of temporal pattern, as well as frequency? Simplicity suggests that some sort of temporal comparison is involved, and the operation of the auditory receptors does introduce an intensity dependent latency of firing onset (inversely proportional to firing rate, described above) that could serve as the cue. This possibility has been noted by several experimenters (for example, Mörchen, 1980) but has received little direct investigation. The most appropriate kind of experiment has only been reported for grasshoppers (von Helverson and von Helverson, 1983): with speakers on each side, they varied the intensity and latency between the sounds, and found

the grasshopper would turn to a small difference in either. It would be interesting to see the same paradigm applied to the cricket: note that the neural coding of different intensities will result in an internal latency difference as well as different firing rates; but latency differences in the sound will not cause a difference in the firing rate; so it should be possible to determine if latency difference alone is sufficient for turning.

How the cricket's neural system determines the difference between the two ears is not explicitly included in the general hypothesis that this difference drives the phonotactic response. Thus Boyd and Lewis (1983) "assuming that the insects use the point of maximum reponse difference between the two ears for localization" or Atkins *et al* (1984) "Directional location of sound is accomplished by turning so that the input from both ears to the central nervous system is balanced" could be taken to be consistent with using either the firing rate or the latency as the measure of response difference. Similarly Weber and Thorson's 'simple' model of "turn towards ear more strongly stimulated" (1988) or Huber and Thorson (1985) "turn toward the ear currently receiving the loudest sound" does not specify what form of encoding of the intensity is used in processing. Nevertheless, the experiments performed do for the most part appear to embody the assumption that firing rate is the critical aspect of the signal: thus Huber (1988) proposes "differences in the firing rates of central auditory neurons are evaluated to determine the phonotactic course" as a non-controversial substitute for the previous statements.

Of course the fact that latency and firing rate are not independent makes it difficult to separate them experimentally. Schmitz *et al* (1982) claim that the shorter latencies (as opposed to higher firing rates) that occur with increased intensity predicts a stronger response at threshold if comparison of latency was the mechanism "Since this prediction is in contrast to our results, we can conclude that *G. campestris* females exploit the bilateral differences in reaction amplitude for detection of sound direction". But it is difficult to see why the relative *difference* in latency should be any smaller than the difference in firing rate between the ears at higher intensities.

Studies that have attempted to manipulate neural firing have shown that lack of

firing on one side will lead to turning to the other side. The most direct test of this was performed by Schildberger and Horner (1988), recording from neurons while the cricket walked on a ball. Hyperpolarization of ON1 caused reduction in turning velocity to an ipsilateral sound; hyperpolarization of AN1 reversed turning so that the animal tracked as though the sound were on the other side. They conclude that the "animal always turns to the side of the more strongly excited AN1": however, this method would equally disturb any latency effects. In less clear-cut results, Atkins *et al* (1984) used selective destruction of neurons in *Acheta domestica* before testing behaviour in an arena: killing ON1 caused errors 45-90° to the intact side except for most attractive syllable period; killing L1 (functionally similar to AN1) caused errors of 45-90° with circling to the intact side in 2 out of 3 animals. Similarly Stout *et al* (1985) found killing L1 (AN1) caused circling and error angle phonotaxis. As yet, the neurobiological methodology does not allow a direct test of the results of altering firing rates independently of altering latency: until this becomes possible only indirect evidence can be obtained.

So far no neural correlates of a comparator have been discovered in the cricket. As discussed above, the reciprocal inhibition of the ON1 neurons does appear to enhance the difference in response, but its effects increase the difference both in firing rate and latency; also, the critical AN1 cell apparently does not receive input from ON1.

Several aspects of the signals that cause successful phonotaxis seem consistent with using latency for comparison rather than firing rate. The signal's low duty cycle (with sound present only about one sixth of the time) does not seem well suited to an independent comparison of firing rate, whereas having an interrupted signal, and hence many onsets, is well suited to latency comparisons. Moreover, it appears that the time between syllable onsets is the most critical parameter in eliciting phontaxis (Thorson *et al*, 1982, see below): varying the syllable length between 10% and 90% of the duty cycle has little effect on phonotactic behaviour, yet substantially alters the number of spikes per syllable, and hence the firing rates (there is no significant adaptation in the response). These considerations do not seem to have been applied to the cricket, although for the grassshopper (where the

female sings and the male approaches) von Helverson and von Helverson (1983) suggest

> "it might make sense biologically for the female song to have such a conspicuous pulse structure: the pulses give the searching male more information about the time shift (produced in his CNS by the conversion of an intensity difference into a latency difference) ... For a grasshopper song to be readily localizable, it should be composed of as many short pulses as possible ..."

### 4.4.3 Low-pass and high-pass filtering

Under the assumption that recognition is an independent process, a number of studies have examined the behavioural and neural response in the cricket to variations in the calling song. Under the alternative hypothesis that 'recognition' is a consequence of using temporal comparison to control taxis, these results can be used as clues for the possible functioning of such a mechanism. The cricket prefers the correct frequency and a certain temporal pattern: what are the features of this preference, and how could they arise in a taxis mechanism?

The frequency specificity is at least partly due to receptor tuning: a higher proportion of the receptors are preferentially tuned to around 5kHz, and the AN1 neuron appears to take input primarily from these neurons, though the width of tuning varies between individual crickets (i.e. it is not sharply tuned). As discussed above, the phase cancellation mechanism also requires specific frequency to be able to generate consistent differences. Together these probably account for frequency preference.

Thorson *et al* (1982) proposed that the syllable repetition rate was a necessary and sufficient temporal cue. Most of their sample of *G.campestris* would track a continuous series of syllables (a 'trill' i.e. without the chirp structure of the natural song) provided the syllables repeated at a rate within 15-50Hz (roughly the range of the syllable rate in the natural song under varied temperature conditions). They note that comparable experiments with *G. bimaculatus* indicate that it may prefer

a chirp structure, whereas *T.oceanus* shows a stronger response to a trill than to its natural song. Doherty *et al* (1985) have shown that some *Acheta domesticus* will not track a trill, and others will track a signal that only has chirp structure. The high degree of interindividual variability, and the range of experimental methods, makes firm conclusions difficult, but it would seem that a mechanism that can track a trill, but requiring only minor modifications to introduce a preference for songs with chirps would be appropriate.

The failure of phonotaxis at fast syllable rates is generally attributed to the time constants of decay in the auditory interneurons (particularly AN1): a rapidly changing song cannot be clearly coded after such low-pass filtering. In fact, the neural and behavioural evidence is a little less clear cut. Schildberger (1985) claims that recognition of the song requires an 7-8ms gap between syllables, and notes smaller gaps don't seem to be coded by AN1. Wohlers and Huber (1982) estimate 'critical pause length' for AN1 to be 15ms, and show a degradation (though not disappearance) of signal copying at 50hz (i.e. with 10ms gaps), but they find the syllable structure equally masked with a 30hz song that has long syllables with short pauses (5ms): such songs *are* successfully tracked by the cricket (Thorson *et al*, 1982). Unfortunately the existance of this inconsistancy has not led to replication or further investigation of the results for different duty cycles.

Hypotheses for slow rate rejection are discussed by Huber and Thorson (1985): they consider a template matching process to be unlikely; a correlation process (using temporal delay and comparison) is argued against by the lack of aliasing effects. However, they don't give a clear mechanism for the high-pass filtering that they suggest. The simplest possibility is summation, with more than one syllable required within a certain time period to get a response. To prevent a continuous signal being summed this would require prior processing such that only onsets contribute to the sum (i.e. an adaptive neuron that reflects changes in the firing rate). Alternatively, if the summation is done on the output of the latency comparison (i.e. neurons that fire when one side is active but not both) then only syllable onsets will be available, because only at onset will one side (the one with shorter latency) be active without the other.

This latter possibility is not generally considered as it is assumed that as recognition takes place independently of location, it involves combining rather than comparing the signals (Huber, 1983, von Helverson and von Helverson, 1987: note however that a straightforward combination of signals with latencies differing by up to 15ms would in fact obscure the 15ms gap in the calling song). Schildberger's (1984) results for brain neurons discussed above (the finding that some fire only to syllable rates within a certain range) are consistent with a summation mechanism (for example, firing doesn't start till after the second syllable occurs), but are not sufficiently detailed to determine exactly what neural connections and processes may be involved; nor was the effect of differences between the ears examined.

### 4.4.4 Comparison not essential

One finding that has received much experimental investigation is that some crickets can still perform phonotaxis when one ear has been removed (by foreleg amputation: locomotion is apparently not significantly disturbed by this). First reported by Huber *et al* (1984), the behaviour was more closely examined on a treadmill by Schmitz *et al* (1988), who described three resulting behaviours: tracking, circling, and 'lagging', i.e., circling but with net movement towards the sound. The result is surprising because under the 'turn to the louder ear' hypothesis, the firing on the intact side should cause continuous turning in that direction.

However it is not necessarily the case that there will be constant response from the intact side. If the main contribution to phase cancellation is from the ipsilateral spiracle, then a single ear's response will vary according to the direction of the sound, and it is possible that it will fall below threshold at certain angles. This may explain at least the 'lagging' behaviour: in circling the animal will pass through an angle where the lack of response on the intact side results in it moving forward until the angle is again sufficient to start a turn. Tracking is more difficult to explain but it has been suggested that the auditory neurons on the amputated side may have some residual firing (possible through reafferentation to the intact

side), which could outweigh that of the intact side at certain angles, leading to a turn the other way.

What this result emphasises is that cricket auditory localization is fundamentally unlike the intensity or phase comparison mechanisms usually discussed for auditory localization (see above) which require two measurements in different locations to be compared centrally. Rather, with a phase-cancellation mechanism, two inputs (the direct and indirect sound waves) are already combined in the response at one tympanum, so neural comparison becomes less essential. On the other hand, with only a single receptor, differences in signal intensity are harder to cope with: the range of intensities over which one-eared phonotaxis occurs appears to be limited in one-eared crickets.

### 4.4.5 No interference

It is clear that crickets show the ability, when presented with multiple attractive sound sources, to track and reach one of those sources without substantial interference: this is frequently the situation occuring in nature. Describing this as 'choosing' a source, however, seems to introduce a non-explanatory cognitive component to the behaviour, as does Weber's (1984) speculation that "females apparently recognise that two attractive patterns are present and that they come from different directions ...".

Huber (1983) reports that on a treadmill, the animal will vacillate (that is alternately track one then the other) between two equal sound sources; but if one is a few decibels louder, it will be consistently followed. Doherty (1985), in arena tests describes a tradeoff between intensity level and temporal characteristics in the song influencing choice: that is, a song with a less typical syllable and chirp rate will be preferred only if substantially louder than the alternative 'correct' song. Stout and McGhee (1988) in similar studies find that varying syllable rate, chirp rate and intensity levels, an intensity difference of 10db will overide other preferences, but for equal intensities, certain combinations of syllable and chirp rates will be preferred. Unfortunately the complex design of these experiments

(not utilising a multivariate design, but rather making an unstructured series of comparisons) make the results difficult to apply to hypotheses about mechanisms.

As the two songs are at the same frequency it is unlikely that auditory processing could separate the input (i.e. recognise 'two songs') as the same neurons will be involved responding. Alternative possibilities are that the combined input nevertheless contains sufficient directional information to follow one source, or that some temporal interaction in the processing means that response to one source blocks response to the others.

In an interesting series of tests, Weber and Thorson (1988) used a choice paradigm on a treadmill where elements of songs were presented alternately from each side. If a normal song was split in this way so that alternate chirps or alternate syllables came from each side, the animal would track a 'target' *between* the two speakers, apparently combining the input from each side. If each side produced a normal song, with the chirps or syllables interleaved (so the combined song would be at double the normal rate), the cricket would vacillate between tracking each source. With a single sound source, altering the intensity of alternate syllables would not disturb tracking: if the alternate syllable came from different sources, the tracking would be biased towards the speaker with *lower* intensity. Weber and Thorson admit the difficulty of interpreting these results because of the lack of knowledge of how a complex sound field will affect the receptor mechanisms. They conclude that the "directions adopted by the females are apparently influenced not only by the 'tracking mechanism' but also by the altered activation of the 'recognizer' as the animal turns in the split-song stimulus field"—perhaps it rather suggests that the operations of taxis and recognition are less separable than is generally imagined.

The idea that the cricket does 'choose' has been reinforced by reports that the females are apparently more attracted to certain males. Zuk (1987) claims females are attracted to older males, though age is not correlated with song intensity or duration. Simmons (1988) finds females attracted to the songs of larger males even when these are intensity balanced, and notes that larger males have a higher syllable rate and (hence) shorter chirps as a possible basis for choice. However,

Bailey *et al* (1990) present evidence that the bush cricket chooses the closer male, rather than larger. The evidence available so far does not seem inconsistant with the idea that the operation of the taxis mechanism is such that the 'better' song, in terms of intensity and recognisability, will effectively dominate the response, rather than selective 'attention' being paid to one song over another.

## 4.5 Proposed mechanism

### 4.5.1 Description

In combination, the considerations and results discussed in the previous section give rise to a hypothesis about the neural structure that could explain phonotaxis in the cricket. The basic principles behind this hypothesis are: that recognition and taxis should not be considered separate processes; that latency comparison is the most effective way to implement taxis that is dependent on a temporal pattern; and that straightforward neural summation properties will underlie the filtering of the signal. It is also suggested that a simple mechanism based on these ideas might in practice suffice to produce some of the more complex aspects of cricket behaviour; that is, more complex processing suggestions shouldn't be introduced until they are proved necessary.

Figure 4–7 shows the hypothesised mechanism. The basic stages are:

- Phase cancellation causes a direction-dependent difference in the effective intensity at the ears: the side closer to the sound will have a stronger response.

- Intensity differences lead to a difference in firing latency in auditory neurons through summation of the input, *i.e.* threshold is reached more quickly for higher intensities.

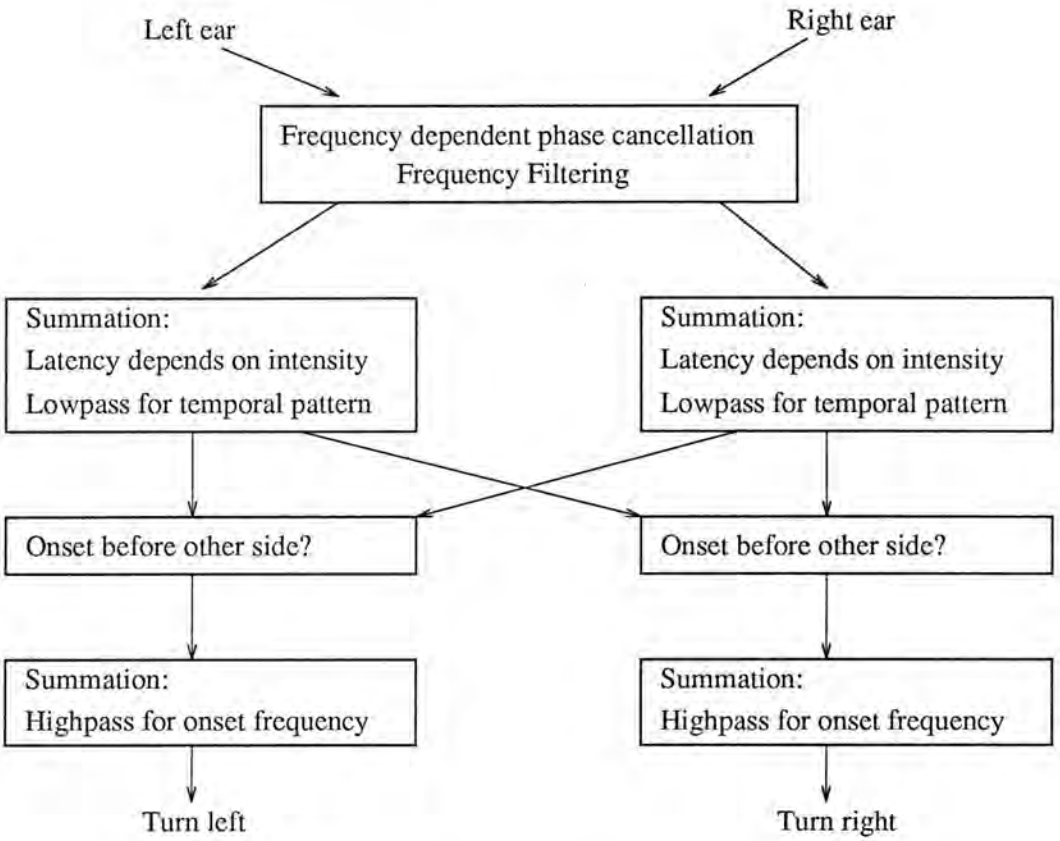



**Figure 4–7:** Hypothesised mechanism for phonotaxis

- Above threshold (firing) level is maintained as long as input is received, and then decays over time. The decay rate is such that small gaps in the signal are missed, thus lowpass filtering occurs.

- The response on one side is compared with the other: if an onset of firing occurs on one side first, a signal for turning is generated. Requiring an onset means that there must be a detectable temporal pattern after the low-pass filtering.

- Turning also involves summation of the input: signals for turning are summed; decay occurs between signals; thus signals must occur sufficiently frequently to initiate a turn; thus high pass filtering of the signal pattern occurs. This summation could be considered a property of the motor system, i.e. turning requires a certain rate of stimulation.

Hence a sound that is closer to one side and has an appropriate temporal pattern will cause a turn to that side. This is quite a simple scheme, but is generally consistent with what is known of the neuroethology of the cricket, and thus a reasonable hypothesis. The aim is to demonstrate whether such a mechanism can produce phonotactic behaviour; and whether the behaviour produced is comparable to that of the cricket.

### 4.5.2 Conclusion

This chapter has involved the description of a particular biological sensorimotor mechanism, and the application of the of the ideas about perceptual systems to derive a hypothesis about the mechanism. Neuroethologists have not produced sufficient explanatory hypotheses for an implementable mechanism to be derived; taking the approach that such a hypothesis is required, a clearer understanding of neuroethological results is possible. This demonstrates the applicability of the first methodological proposal given in Chapter 3.

In the remainder of this thesis I attempt to establish the viability of this hypothesis by implementing it in a robot model. In doing so, the second and third methodological proposals—using robots and adequate evaluation methods—will be demonstrated.

# Chapter 5

# Robot Phonotaxis: Methods

## 5.1 Aims

The aim of the work described in the next two chapters was to test hypotheses about cricket phonotaxis by building a robot that could perform phonotaxis, *i.e.* move towards and arrive at a sound source. Thus, the mechanism embodied in the robot was to be a model of the hypothesized mechanism of phonotaxis in the cricket. Therefore there needed to be an adequate representation of the problem, and this guided a number of decisions about what was built.

In chapter 2, I argued that the nature of the signal significantly affects the mechanisms of sensory motor control. Hence it was specifically sound that was to be used as the medium for taxis, rather than a substitute physical signal (it is not uncommon in robotics to investigate, for example, 'visual mapping' using sonar signals; and in simulation, signals are often modelled as a generic 'sensory source'). In the next section some pilot experimentation with phototaxis is described: the problem posed is sufficiently different to make it difficult to examine the relevant hypotheses.

The robustness of phonotaxis in apparently adverse conditions is also one of its major features. To reflect this, and avoid the assumptions inherent in purely computational simulations (where noise and uncertainty has to be specifically included) it was important to use a physical robot in a real environment. For practical reasons a small, cheap robot is preferable — demonstrating a viable mechanism under such conditions not only points to the robustness of the mechanism but also

to the viability of using relatively simple robots to explore neuroethological problems. A consequence is that precise motion control and highly accurate sensors are not available and the mechanism really must deal with real-world uncertainty. However, small, cheap six-legged robots with capability for sensory extensions are not readily available, and building one was beyond the scope of this thesis, so the robot used was wheel-based: the possible affects of this motor difference are discussed in several places below.

Given wheels, the environment was also necessarily different to the uneven terrain of the cricket's natural environment, although perhaps more comparable to conditions in experimental paradigms, with the basically smooth floor of the laboratory as the substrate. Enviromental complexity was further limited (for most of the experiments) by the requirements of the methods of evaluation (see 5.7) with the robot moving only inside an enclosure to keep it within view of an overhead camera, rather than encountering the full range of features in the lab. However, no attempt was made to make the sound field more ideal, with the effects of reflection etc. being left to occur; and background noises uncontrolled.

It is an essential part of cricket phonotactic behaviour that it only approaches a certain sound signal. In the previous chapter I have argued that this is characteristic of the functioning of the mechanism as well as of the observed behaviour. Thus a sucessful robot would approach a sound of specific frequency and temporal pattern, and ignore other sound sources. It was hoped that a number other features of cricket behaviour might also be reproduced by the robot, thus demonstrating that the hypothesised mechanism, implemented in the robot, may explain them. These specific tests include characteristics of the path, recognition effects, behaviour with obstacles and two sound sources: the experiments and results are described in detail in the next chapter.

## 5.2 Previous work

The large amount of published research in the field of mobile robotics contains very few examples of robots performing taxis (following a sensory gradient to find its source), especially to sound sources. Although the problem of getting to a specific location is frequently addressed, in the vast majority of cases it is approached by attempting to model the environment and plan a path through it. This 'general navigation' problem has been much studied but as yet yielded few really successful systems. It is somewhat surprising that something as useful as finding the way towards specific signals has received so little investigation.

### 5.2.1 Robots doing taxis

Probably the first example of robot taxis was one of the first examples of robotics: Grey Walter's 'turtles' (1953). These devices were hard-wired to approach a light source (the location of the recharger), having a front 'steering' wheel that would alter direction unless a certain threshold level of light was received by a sensor pointing that direction. Thus it would circle until a light was detected and then move towards it. However, if it lost the light, the mechanism did not allow it to turn back towards it directly: it would circle again till the light was found. In fact this behaviour is rather reminiscent of the 'lagging' phonotaxis of one-eared cricket, and shows how a single directional detector can be sufficient for finding a sensory source.

Other more recent mobile robots have been programmed with a tendency to orient themselves in a sensory field, often to demonstrate their turning mechanisms rather than to investigate the efficacy of getting about this way: for example, Brooks' robot *Genghis* (1989) turns towards infra-red (heat) and thus will follow a (slow-moving) human. It will also find its way into a flame thrower, which perhaps demonstrates the utility of taxis that is selective for certain signals! Tracking mechanisms (ranging from Connell's (1987) *Tom and Jerry* to mobile cameras)

are in some sense implementing similar behaviour; although perhaps a distinction should be drawn between orientation based on having many sensors (or pixels) pointing in different directions, and orientation based on one or two sensors that respond over a range of directions—it is difficult to conceptualise a visual point as constituting a sensory *gradient*.

A behaviour akin to taxis is provided in systems that make use of 'potential fields' to direct the movement of the robot. In this case the gradient is not directly present in the signal energy, but is created in the sensory processes of the robot. The subsequent effect is a taxis-like behavioural response—turning according to the relative strength ('potential') in various directions. However, such methods usually involve multiple repellant sources rather than a single attractive one (or a combination of both, e.g. Arkin (1989)), thus the situation is different to the problem of locating a particular attractive signal source. Payton (1991) describes a system that calculates a gradient field based on routes to a goal, and points out several advantages of using gradient-based location rather than plans: behaviour can be reflexive (the movement decision based only on current information); errors can be recovered; and unexpected opportunities taken. However he does not report a robotic implementation of this scheme.

### 5.2.2 Lego phototaxis

One of the sensory inputs developed for the Lego vehicle technology used in this work (see below) was based on light dependent resistors, which could be used to get the robot to approach a bright light source. While developing the phonotaxis system some experiments were done with this, and it proved quite easy to get a robot to find a light, even with only a binary response on each side. This ease was due mostly to the straight line travel of light: if a sensor is turned away from the source, it will receive effectively no light. Hence whatever the distance or intensity of that source, it is easy to distinguish a sensor pointed towards it from one pointed away, and to turn appropriately.

Other research with the same robot equipment also used phototaxis as one of the tasks for robots while investigating action selection or agent co-operation (Pebody, 1991; Mein, 1991). In this case there were several light sources, but the robots had little difficulty in approaching the nearest light.

### 5.2.3 Phonotaxis

Although auditory information is used as orientation information in a number of robot systems (for example, Donnett, 1993; or as sonar, e.g. Kriegman *et al*, 1987), my attempts to find examples of implemented taxis towards sound sources were unsuccessful. Flynn and Brooks' (1989) *Squirt*, designed as a demonstration of how small a robot could be, used two ears and detected amplitude differences to turn in the direction of sounds, but this aspect of its behaviour wasn't particularly successful.

## 5.3 Construction

### 5.3.1 The lego robot

The robot used here is based on the Lego vehicle technology developed by Donnett and Smithers (1990), which uses Lego Technic (car kit no. 8865) as the physical base and adds electronic circuits for sensor devices, motor interface and behaviour controllers. This technology has proved adaptable for a range of experiments on robot behaviour (Dallas, 1990; Pebody, 1991; Mein, 1991; see also descriptions of the use of a similar Lego based robot system developed by Resnick (1991)). In particular, it allows custom design and easy redesign of the body, as well as flexibility in the addition of alternative circuitry. The work described here is perhaps the most complex behavioural task yet attempted using such robots, and the current robot represents the accretion of design experience from attempts to implement a variety of behaviours.

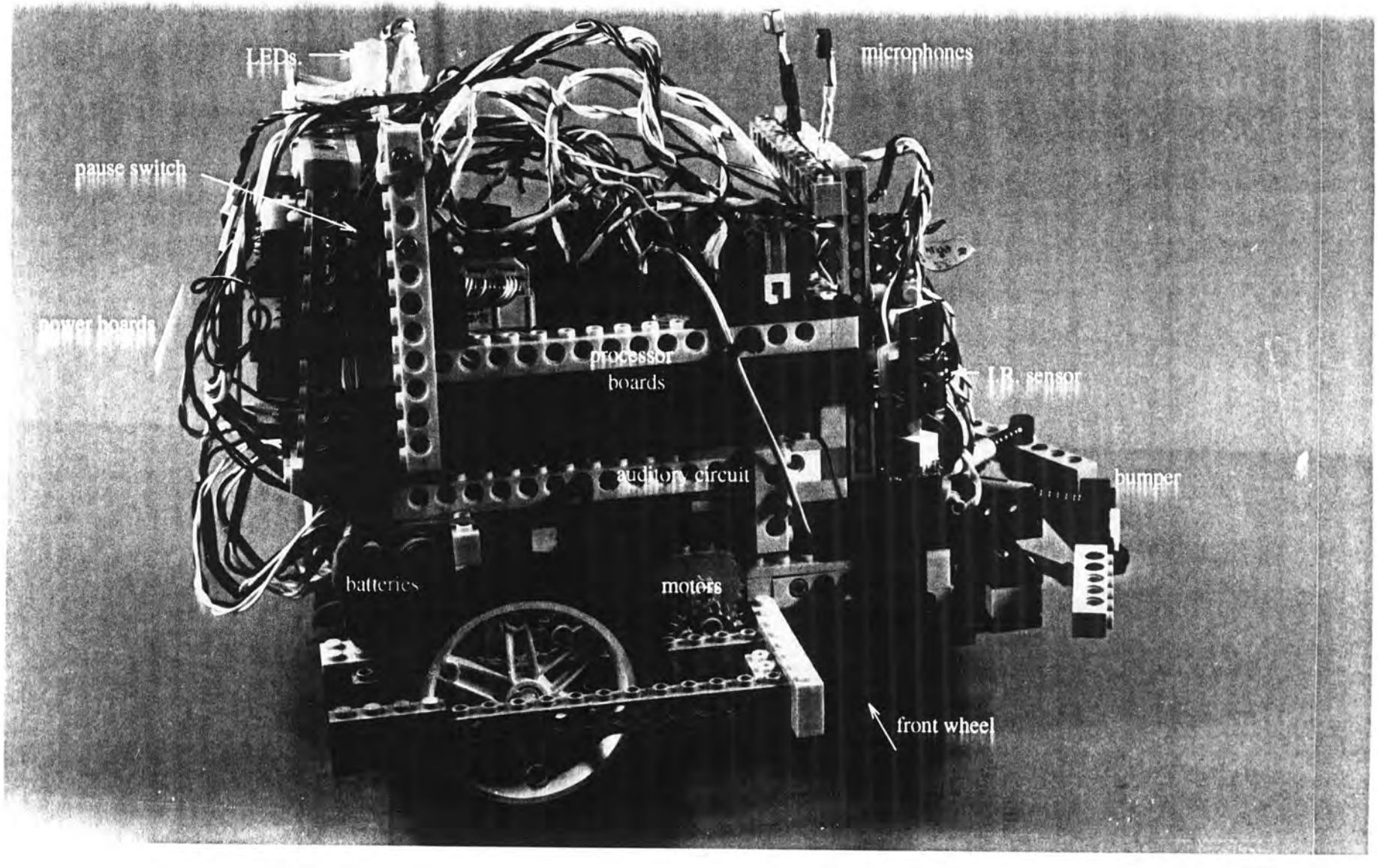

LEDs
microphones
pause switch
power boards
processor boards
I.R. sensor
auditory circuit
bumper
batteries
motors
front wheel

Figure 5–1 is a photograph of the robot.

**Dimensions:** The robot is 19cm high, 21cm wide, 27cm long, and weighs 1350g with an additional 350g for the sets of batteries.

**Frame:** The upper part of the robot supports the circuit boards: two 10x12cm processor boards, two 8x12cm auditory boards, two 8x4.5cm power boards and an 8x4.5cm infra-red (IR) board. The frame design attempted to minimise the space occupied by these boards and the amount of additional weight used structurally. Start/Stop and Pause switches are attached to the back, with a set of 4 LEDs that are used for tracking the movement of the robot.

**Chassis:** The lower part of the robot contains the motors, gears and wheels. Two 4.5V Lego motors are connected by five gear wheels (total gear ratio is 25:1) to 8cm diameter wheels, with 0.5cm wide tyres. The design is intended to improve the stiffness of the axles without impeding free gear and wheel movement. The front 'wheel' is actually a fixed ball-bearing, which on a smooth floor offers minimal resistance to movement in any direction, allowing immediate response to direction changes determined by the controlled movement of the rear wheels.

**Power supply:** two sets of 8 rechargable Nickel-Cadmium cells (AA), providing a total of 4.8Wh each. One set power the processor and auditory circuits, the other powers the IR circuit and the motors. This eliminated previous interference coupled through the power supply from the IR to the auditory circuits, and prevented current pulses drawn by the motors resetting the processor. Fully charged cells provide up to 20 minutes operation, though the robot was rarely run continuously for that length of time.

**Movement:** The power provided to the motors is controlled by signals sent from the processor. They can be driven forwards or backwards using either all or three-quarters of the available battery power (in practice the slower speed was not used in this robot), or stopped. On full forward the robot would

accelerate over 50cm to a top speed of around 50cm/second, although this speed depends on the level of charge in the batteries. Turning, in this implemention, is done by stopping the inside wheel (faster turns could be made by reversing the wheel). The battery power is considered too low when it is no longer sufficient to execute perceptible turns within 500ms.

**Obstacle sensing:** a hinge-mounted bumper is attached to the front of the vehicle. This rocks left or right depending on the position of impact, and thus presses one of two microswitches so that collision with obstacles can be detected. To avoid obstacles before collision where possible, there are also two infra-red sensors, emitting a signal that is detected if reflected by a nearby object. The detection is binary with the level set so as to respond at a distance of about 30cm from the wooden walls used as barriers in the experiments described below. Each detects obstacles over roughly a 60° angle, and they are directed 30° to each side of the midline.

### 5.3.2 Auditory Circuit

To implement a cricket-like processing of the sound it was necessary to devise a circuit that could process the phase and amplitude of sound signals received from two microphones (lacking the means to build a physical phase cancellation device, dedicated electronics provide sufficient speed to detect phase differences). The initial design was intended to provide a reading for each syllable in the signal, by using low pass filtering to detect the onset and offset and to use these to start and stop the reading. However it was found that the level of noise and reflection made it impossible to get a sufficiently clean signal for this purpose, so the circuit was modified to allow the processor to initiate and reset readings, which occured every 32ms. Noise problems were particularly bad for the low frequency (500Hz) originally chosen for a comparable wavelength:size ratio to the cricket, and so a frequency of 2kHz was adopted as a compromise.

A schematic diagram of the circuit is given in Figure 5–2; the circuit diagram is contained in Appendix A.

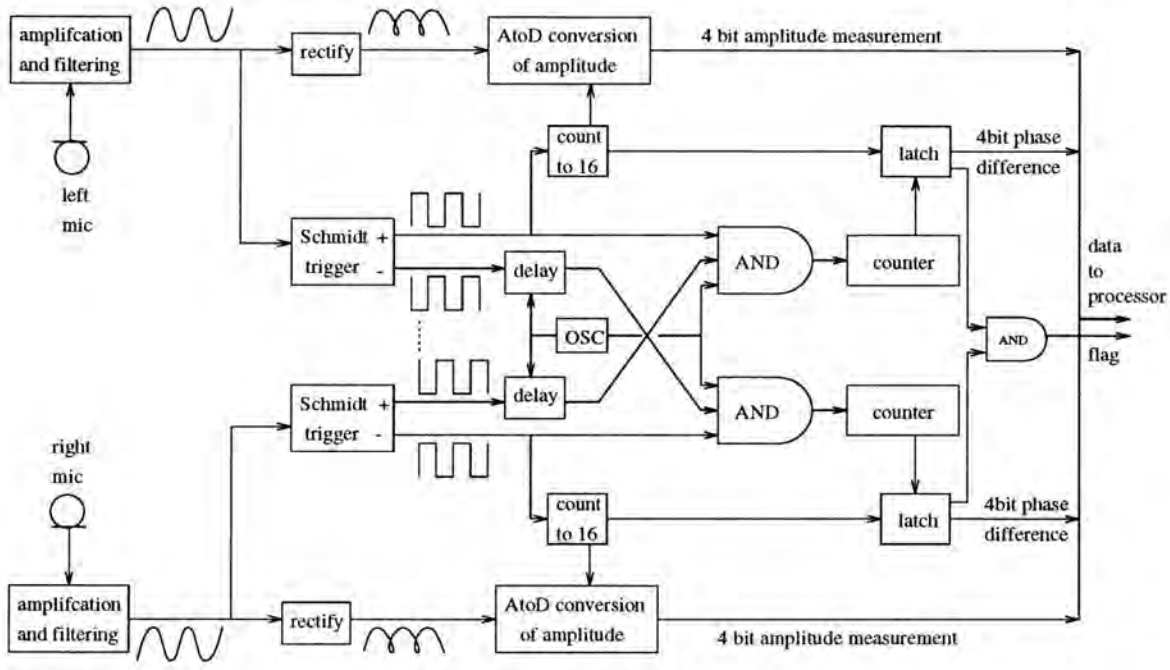



**Figure 5–2:** Schematic of auditory circuit

**Microphones and initial filtering:** The two microphones mounted on top of the robot are Knowles miniature microphones. They are positioned 4cm above the frame and set 4.5cm apart (aproximately 1/4 the wavelength at 2kHz), symmetric about the midline. They face backwards as this was found to produce a cleaner phase response than other positions (it is interesting to note that the cricket's main auditory input is the posterior, *i.e* rear-facing, tympani). For each microphone, the signal is amplified and then filtered with a bandpass filter centred on 2kHz. It then is fed simultaneously into the amplitude and phase readers.

**Phase:** The amplified signal passes through a Schmidt trigger to create a square wave. The inverted output passes through a shift register which creates a 125$\mu$s delay (1/4 the period of 2kHz). The delayed signal is ANDed with the undelayed, uninverted signal from the opposite side: the output of this comparison gates the input of a 1024kHz oscillation to increment a counter. After 64 cycles of 2kHz sound, the high bits of the counter are latched, giving a reading of 0-15 for the phase, and setting a 'data ready' flag. As an

example of this function, consider the vehicle to be perpendicular to the sound. The near side microphone should have a phase 90° in advance of the far side microphone, so the delayed, inverted signal will cancel out the farside response; correspondingly the far side, 90° behind and delayed a further 90° will, when inverted, coincide with the nearside signal. In general, the difference between the external phase (determined by orientation in the sound field) and the internal phase delay alters the overlap of the compared signals and thus the contribution to the counter. This is the same principle as the phase cancellation in the cricket, although in fact, ignoring amplitude differences, it does not represent the tympanum response but only the effective phase difference: amplitude and phase readings were combined into a 'tympanum response' value in the program (see below).

**Amplitude:** the signal is rectified and stored in a capacitor: when the 64 cycles latch the phase, the stored value is fed to an analog to digital converter which provides a four-bit estimate of the amplitude of the sound.

The readings provided by this circuit are only accurate if the sound is continuous during the 32ms readings, which cannot be guaranteed with a temporally patterned signal. The processor handles this by initiating a reading then 35ms later looking for the flag signalling latched data. If data has not been latched, less than 64 cycles have occured within the time period, and a zero reading is recorded and the counters reset (see Figure 5–5). Note that both this and the 125$\mu$s delay are determined by the frequency; signals at other frequencies will not give accurate results.

Readings from the circuit are shown for two places in the room (Figure 5–3), with the robot facing in eight different directions. One metre from the speaker, the values for left and right phase clearly correspond to the appropriate direction; at three metres reflected sound causes substantial distortion although the directionality is still approximately preserved. The amplitude measurements show no clear directional effects, but are reduced by distance, as would be expected. There was substantial variability at any one location, and the change in phase and amplitude

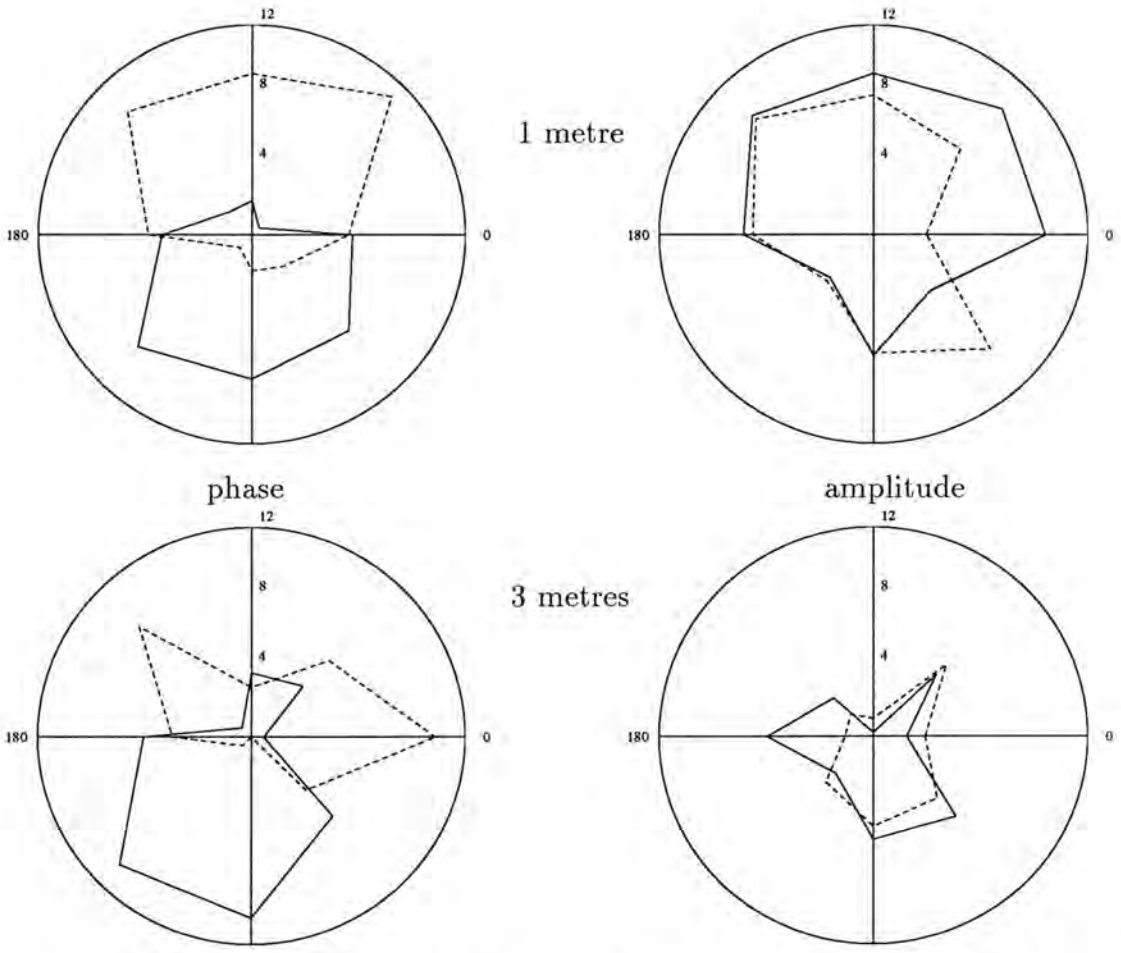



**Figure 5–3:** Four-bit values (averaged over 50 circuit readings), of phase and amplitude, at different distances, when the robot faces different directions (with the speaker at 0°). Solid line is left ear, dotted line right ear.

across the room was not particularly consistant with the ideal. There was also some noise generated by the motors and movement, though in general this didn't obscure the signal. Phase readings for a temporally patterned signal are shown in Figure 5–10.

### 5.3.3 Processor

The processor is a PSI Microsystems single board computer using a 68000 clone processor running at 30MHz, with a custom-built input/output board compatible with the other circuits. Programs are written in C and compiled on a Sun3; the compiled program is transferred to a Viglen PC from where it can be downloaded

to the robot processor via a serial communication link, which can be detached to allow autonomous operation. The processor uses 16 output ports, to control the motors, signals to the auditory circuit and display lights; and 16 input ports, that take data from the the start and pause switches, the bumper and IR sensors and the auditory circuit. Configuration of the processor is described in Colles (1992).

The processor supports pseudo-parallel processing: a set of tasks are cycled sequentially through the processor at a rate of 250Hz (i.e. each gets 4ms processing before being put to the back of the queue). Tasks can be forked, suspended or killed; the control most frequently used is rescheduling, which puts the current task to the back of the queue immediately, pre-empting the elapse of 4ms processing time. This allows efficient use of multiple tasks that mostly watch for a certain condition and only then do more complex processing; several such condition tests can be made by separate tasks in less than 4ms.

## 5.4 Robot interface

The majority of the code written for this project is concerned with the interface to the robot's sensors and controls. The characteristics and functions described here were important in establishing this interface; and show the influence of the hardware on the control program that implements the phonotaxis hypothesis (see section 5.6).

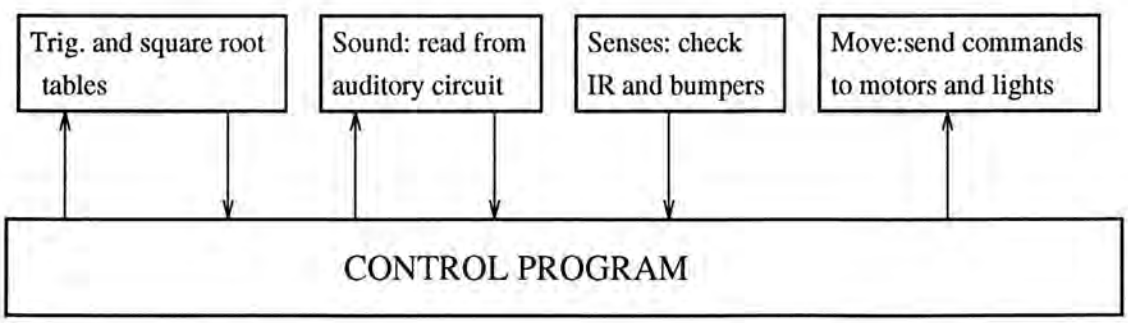



**Figure 5–4:** Functions for interface to robot.

**Speed and calculation:** The fact that the auditory circuit provides information every 32ms means that processing speed is quite critical. It was sometimes difficult to ignore the fact that processing was not actually parallel. Also, the complexity of calculations had to be minimised: only integer arithmetic was used and bit-wise operations (such as shifting for division) were often employed to speed it up. Some trigonometric values were required: these were stored in an array and accessed as a lookup table. In general there were strong constraints on the detail with which the proposed system could be implemented: for example, neural functions were represented very simply.

**Motor commands:** The motor control consists of six bits specifying direction and power for each motor, plus a seventh enable bit. Movement commands in the program (to start or stop each motor) change the value of variables *rightstate* and *leftstate* which are combined to form the current *motorstate*. Pauses are implemented by disenabling, so that when a pause ends, the type of movement prior to the pause will be restored. The eighth bit is calculated from the motor states to reflect whether the vehicle was moving forward or turning, and this switches two of the large LEDs, signalling the state.

**Sensor readings:** Bumper and IR sensing is quite simple—checking the appropriate bit of input, which changes when the sensor state changes. For the auditory input, (see Figure 5-5) initiating a reading from the control program sends a signal to the auditory circuit to stop the counters, and check for the flag. If the flag is set, the latched data is read (the 8 bits for the left side followed by the 8 bits for the right), otherwise a zero reading is recorded. The values of the readings, the time, and whether the current motorstate is turning or forward (the eighth bit above, this allows internal and external data to be correlated), are stored in an array, then the counters are reset and restarted.

**Task generation and control:** The MAIN task in the processor controls the other processes. When the start switch is pressed it signals with the LEDs and then forks the other tasks (which constitute the control program). These

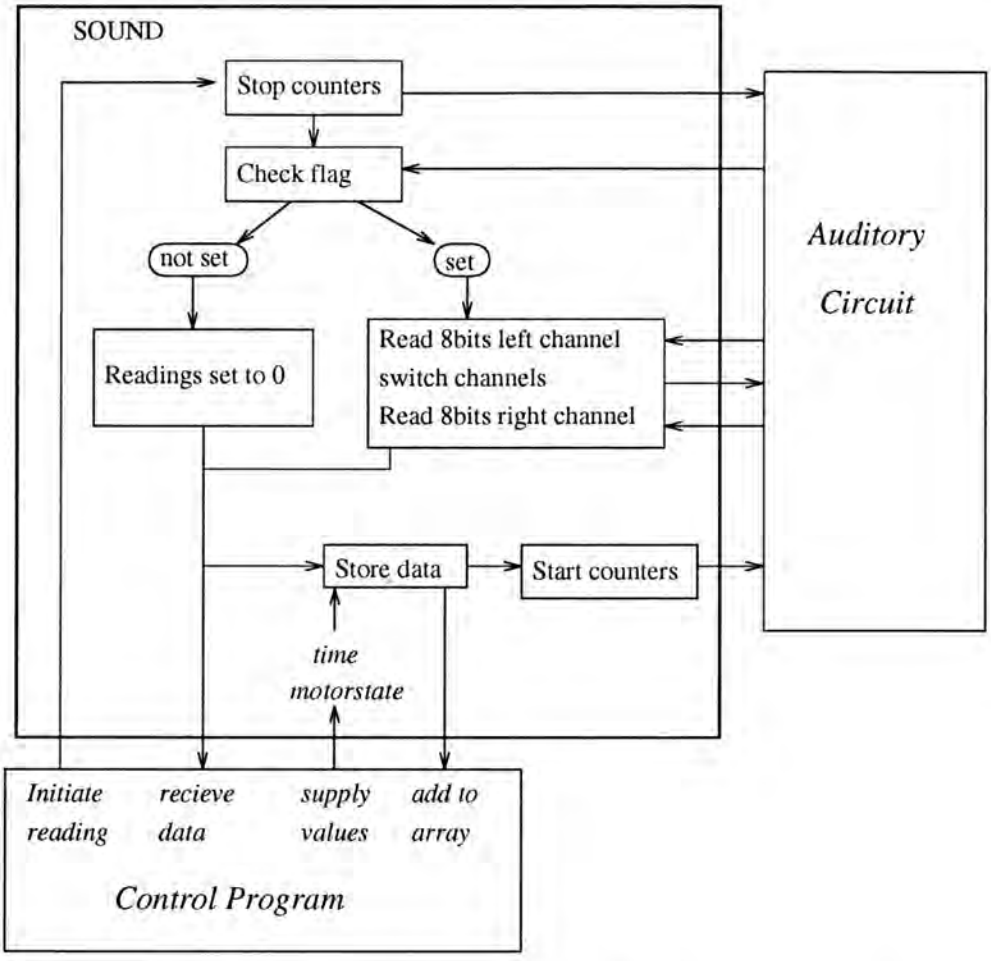



**Figure 5–5:** Operation of the 'sound' function interface to the auditory circuit.

can be suspended and restarted using the pause switch, or killed (at the end of a trial) using the start switch, or by the index of the array used to store data exceeding 2000. When the tasks are dead (and the robot thus stopped) it can be reconnected to the serial port on the PC and the contents of the array downloaded. The processes can be restarted for the next trial.

## 5.5 Simulation Interface

As the auditory circuit took some time to be built, the viability of a simulation interface for initial tests of the control program was explored. An alternative set of functions was devised, equivalent to those described above, but instead of sending and receiving messages from motors and sensors, they would estimate the

movement and simulate the sensory response. Thus the same control program could be run, interacting either with the robot or with the simulation.

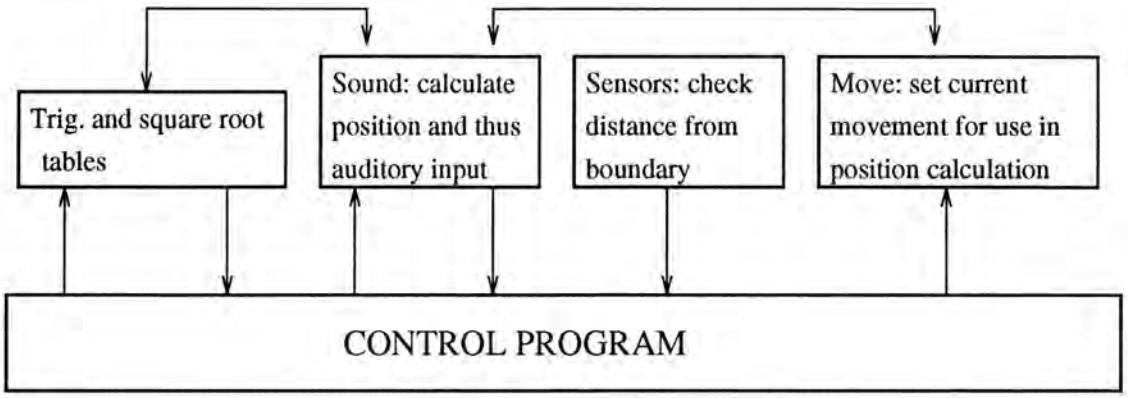



**Figure 5–6:** Functions to simulate robot behaviour.

At each signal for a sound reading (every 35ms) the simulation calculates the change in position (polar coordinates centred on the sound source, i.e. the distance $d$ and angle $\alpha$; plus the forward direction $\theta$ relative to $\alpha$, see Figure 5–7) based on the current motor state. This requires estimates of forward movement speed and turning speed: to be accurate it would need more complex dynamics such as accelaration and battery level dependence, but even at this level the calculation of the new position took a significant amount of time relative to the operating speed of the rest of the program (around 2 seconds). To maintain the time dependent factors in the control program, all the other tasks are suspended while this calculation takes place. The phase and amplitude for the new location are then estimated: these are based on an ideal point-source sound emitter, with amplitude decreasing as the square of distance, and phase difference based on the cosine of relative direction $\theta$, ignoring possible reflection or interference. The temporal patterning was defined by time variables: if the current time fell between syllables, the auditory values were zero. The bump and IR senses were modeled as 90° obstacle detection on each side, bump 10cm and IR 40cm in front of the robot's position, with the simplest boundary being a circle at a radius of 4 metres. (the equations used for all the calculations are in the program code in Appendix B).

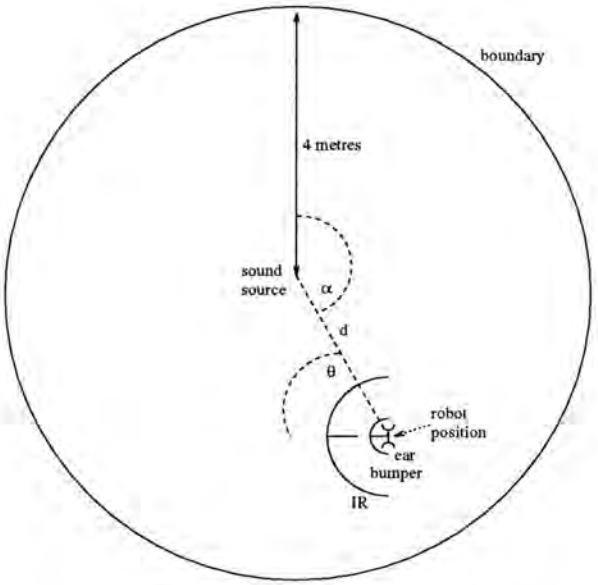



**Figure 5–7:** Simulation environment

The complexity of this simulation (in terms of depth of modelling of sensory and motor interaction with the world) compares favourably with many used in investigations of 'autonomous agents', such as those described in Chapter 3. However it was of limited use in designing the control program:

• With perfect input data and exact movements, it was hard to think of a program that *wouldn't* turn the simulated robot more or less directly towards the speaker and take it in a straight run towards it.

• The number of parameters that required estimation for the simulation was far larger than the number in the control program itself, but without accurate estimation (which required a working robot to measure from) the results were of little use for tuning the control program. For example, the simulation could easily run with a rapid syllable rate, such as 35ms on, 35ms off, giving one measurement per syllable followed by a zero for the gap. But this would suggest different means to create the latency difference, to detect gaps, and to have the net effect on movement, than those that were eventually employed in the program when operating off the necessarily longer syllables required in the real situation.

• Extending the simulation to include noise or the other more complex environmental situations used in testing the robot (see Chapter 6) would have required far

more extensive modeling of the acoustic field and obstacle detection, the difficulty of which would probably outweigh the benefit.

Attempts to improve the usefulness of the simulation were abandoned once the auditory circuit became available. There is further discussion of problems with simulation for this task in Chapter 7.

## 5.6 Control Program

The control program was designed to embody the hypothesis about the mechanism for cricket phonotaxis discussed in the previous chapter. Within that requirement it was kept as simple as possible; that is, with few functional elements other than those explicitly mentioned in the hypothesis. It is undoubtably a simplification of any actual process in the cricket: the idea is to include only those elements that are supposed to be required for phonotaxis to see if they are indeed sufficient to explain the main aspects of the behaviour found in the cricket. Rather than detailed neural modeling, only the essential dynamics suggested by the mechanism are represented. This as much as anything is dictated by the necessity for fast calculation for real-time control (see above).

Excluding the MAIN module described in 5.4, the behaviour of the robot is controlled by five concurrent processes that can be considered as corresponding to three groups of neurons in the cricket: the ascending auditory processing; the central brain comparison; and the descending motor control. Although the following description presents them in order, as each controls variables that affect the behaviour of the next, it is misleading to see their functions as a sequence of processing steps as their temporal interdependence is a critical feature of their operation. Figure 5–8 shows the interconnection. The code is in Appendix B.

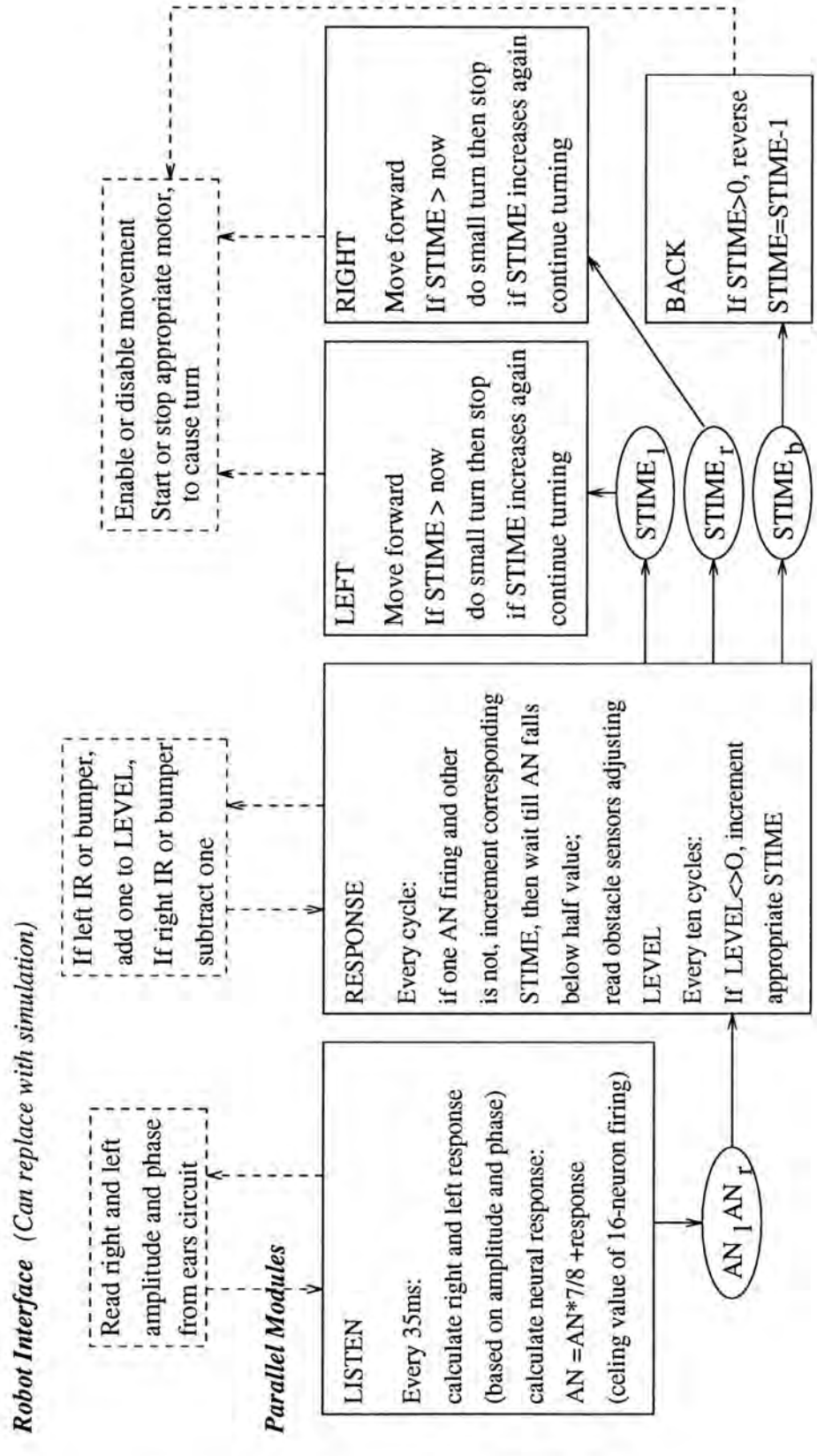



**Figure 5–8:** The control program.

### 5.6.1 The LISTEN module

This module controls the auditory circuit, as described above, looking for amplitude and phase values every 35ms and inititating the next reading. While a reading is taking place in the circuit, the previous input is being processed: the response equivalent to tympanal movement on each side is calculated as

$$l_{response} = \sqrt{l_{amp}^2 + r_{amp}^2 - 2l_{amp}r_{amp}cos(12 * l_{ph})}$$
$$r_{response} = \sqrt{l_{amp}^2 + r_{amp}^2 - 2l_{amp}r_{amp}cos(12 * r_{ph})}$$

(Note that the phase readings $l_{ph}$ and $r_{ph}$ are four bit values (from the circuit) and thus range from 0—15, so multiplication by 12 gives a range of 0—180°. Similarly, the amplitude readings, $l_{amp}$ and $r_{amp}$, range from 0—15, so the *response* values will be between 0 and 30.) This is the basic model for the cancellation of waves of the same frequency, different amplitude and a particular relative phase, arriving from opposite directions. As discussed in Chapter 4, this is simpler than the actual interaction of waves in the cricket trachea, but no alternative model has been shown to provide a significantly more accurate estimate of the effective amplitude difference caused at the tympanum by cancellation.

The 'neural' response to this 'tympanal' response is represented in the variables $an_L$ and $an_R$ which combine the current response with their previous value using:

$$an = \frac{an * 7}{8} + response,\ if\ an \geq 16,\ an = 16$$

This equation embodies three aspects of the response of the auditory interneurons in the cricket that are considered relevant to the mechanism implemented here, if the ceiling value 16 of $an$ is considered to represent the point at which the neuron starts firing. Recall from Chapter 4 that the auditory neurons in the cricket have response properties characterised by overall increase in in post-synaptic potential with superimposed spikes: the ceiling value of $an$ here represents the onset and duration of the spiking but not the rate or number of spikes.

First, different response values will result in different rates of increase of $an$ towards 16. In other words, the latency till firing commences will vary with the response level (see Figure 5–9). Second, the response is maintained rather than adaptive; the neuron continues to fire while there is a response in the tympanum.

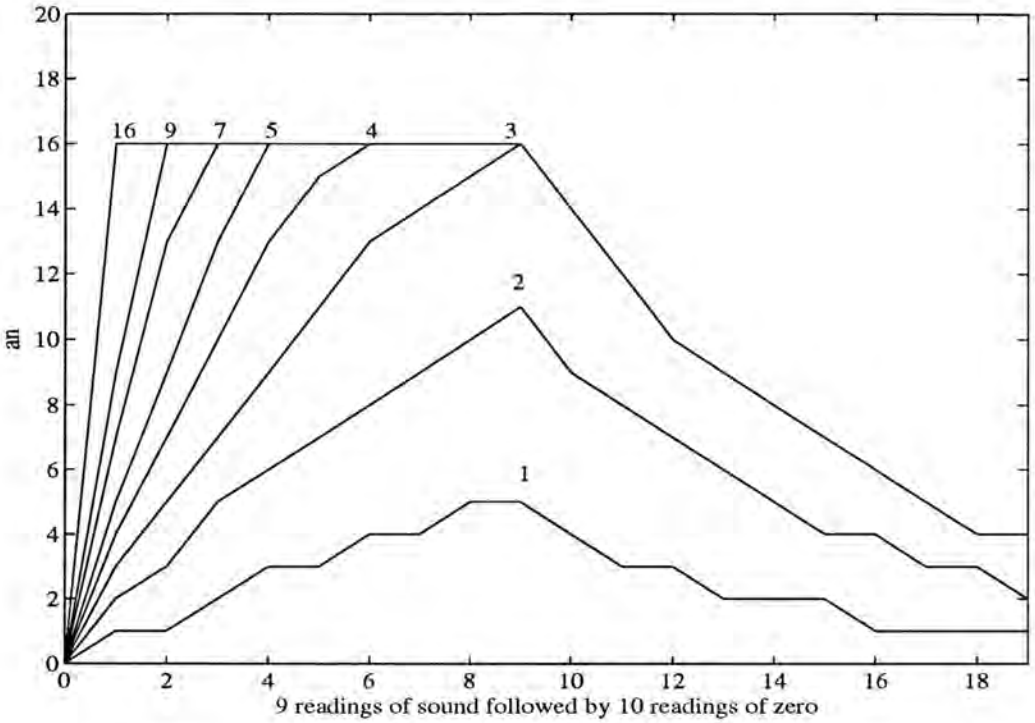



**Figure 5–9:** Latency till *an* reaches 16 for *response* values 1-16.

Third, when the response ceases *an* shows a gradual decrease towards zero at a fixed rate: if the next response occurs before this decay period has elapsed, the rise time to 16 will be more rapid. Thus fast variation in the temporal pattern will tend to become obscured in the *an* response, which is essentially a moving average.

The value 7/8 was chosen to give the best difference in latency for typical response values (it can be shown that with a phase difference of more than two between the ears, there will be a latency difference in *an* for any amplitude, see Appendix C), and also to give a decrease rate that allowed viable onset detection. This interactively determined the 'ideal' syllable rate, where a low amplitude syllable would still reach threshold (*i.e.* last for 9 readings — 9x35ms = 265ms) and the gap would be sufficient to return from threshold towards zero (*i.e.* 10 readings — 10x35 =350ms), in other words an intersyllable period of about 600ms, a rate of 1.6Hz. In practice signals were much noisier than expected, but this filtering was still adequate: Figure 5–10 shows the phase, *response* and *an* values over 10 syllables (6 seconds) for one of the start positions of the robot, with the detectable 'prior onsets' (see below) indicated.

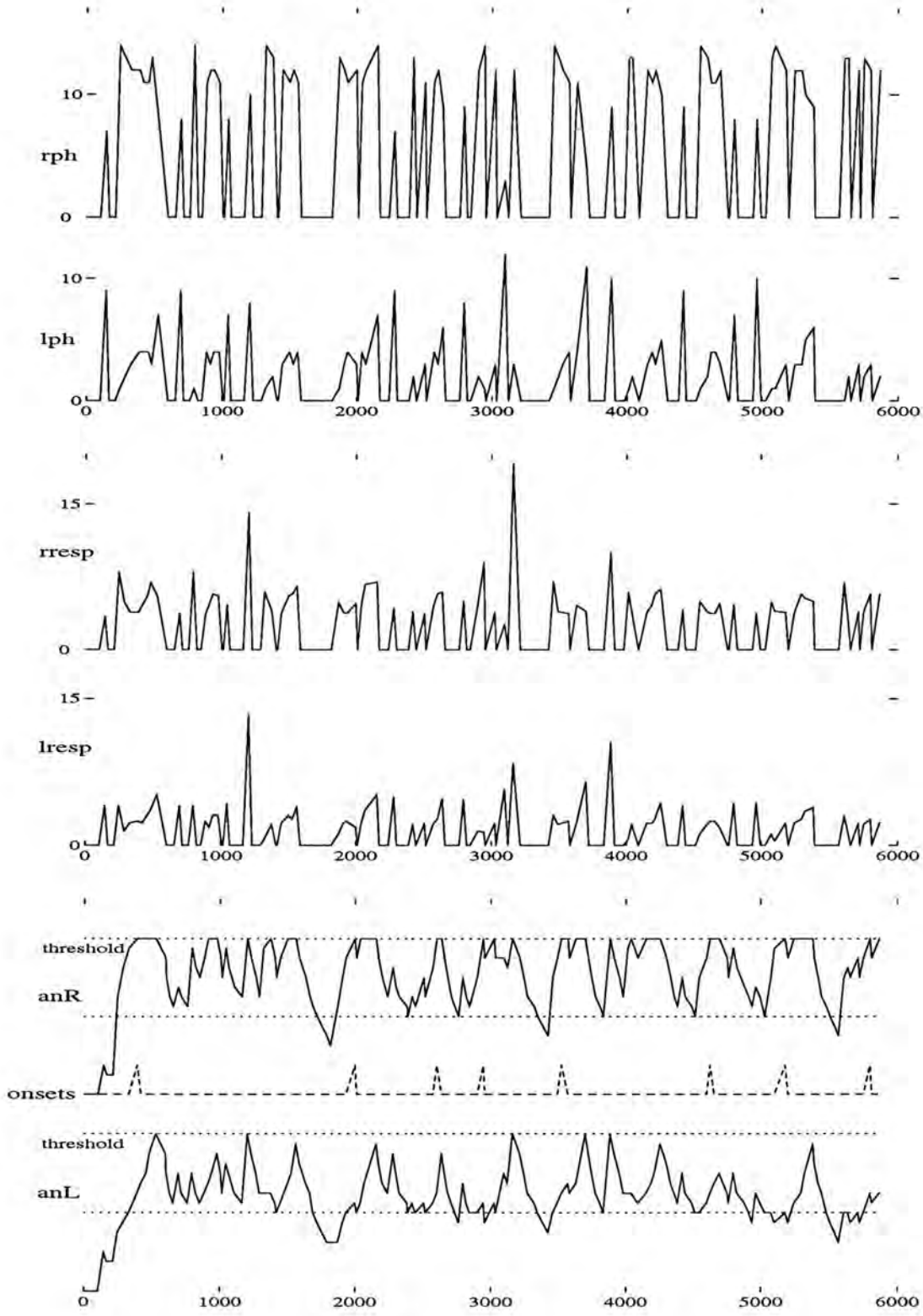



**Figure 5–10:** Actual phase values, *response* values and *an* values with a patterned signal coming from the right: onsets are when $an_R$ goes from half-value to reach threshold before $an_L$

### 5.6.2 The RESPONSE module

This module produces a response based on the value of sensory data. The response is encoded in three time variables, $stime_L$, $stime_R$ and $stime_B$ which specify the 'duration of firing' of central signals for a left turn, right turn or reverse respectively. $stime_B$ is increased by bumper contact (+1 for each cycle that a bumper switch is on); increases in $stime_L$ or $stime_R$ are more complex, depending on both obstacle and auditory sensing:

Obstacle sensing: The response of the IR and bumper switches is sampled over 10 cycles (about 150ms, depending on rescheduling of other tasks) during which the variable *level* is increased by 1 for signals from the right, and decreased by 1 for signals from the left. A net increase from the sample causes $stime_L$ to be incremented (signals from the right cause a turn to the left to avoid the obstacle), a net decrease causes $stime_R$ to be incremented. The constant increment was determined by the motor modules as explained below: the sample size was chosen so that the increment would occur with sufficient frequency yet the sample adequately remove most noise from the IR sensors. Note that identical response from both IRs would cancel out and lead to no response: but the nature of the IR sensors and the positioning of the sensors on the robot meant that in practice this almost never occurred: if it did the resulting collision would cause the required reverse.

Auditory sensing: The $an_L$ and $an_R$ responses determined by the LISTEN module were continuously compared. If one reached 16 before the other, the corresponding *stime* variable would be incremented (a faster onset from the left would cause turning to the left to approach the sound source). Thus a difference in latency was the variable for comparison between the auditory responses, the comparison occuring through simple priority. Further comparison was 'inhibited' until the *an* variable had dropped to half the value.

Latency comparison requires an identifiable temporal event for which signal onset is the most obvious. This means that the presence of onsets is necessary for a response. This requirement, combined with the moving average effect of *an* calculation, constitutes a low-pass filter for the rate of temporal alteration of the signal.

### 5.6.3 The LEFT, RIGHT AND BACKUP modules

These three modules together control the actual motor response to the *stime* variables. BACKUP is the simplest: when $stime_B$ exceeds zero, the motors are put into reverse and $stime_B$ is decremented by one for each cycle through the modules until it is zero and the robot resumes forward movement. The reverse allows it to clear the obstacle (turns generally don't work when the robot is wedged against something): notice that the bumper signals will have also contributed to either the $stime_L$ or $stime_R$ responses thus causing the appropriate turn to follow the reverse.

In the first version of the control program used in the initial tests below, the LEFT and RIGHT modules controlled the left and right motors independently, i.e. stopping and starting that motor on the basis of the respective *stime* value without reference to what the other motor was doing. Each module would execute the following loop:

```
set the motor going forward;
wait until stime is increased;
then stop the motor, causing a turn in that direction;
after a short time (small turn) halt all movement
    by disabling motor response;
wait until stime has elapsed:
    if stime is increased while waiting,
    re-enable motors so that turning continues;
then go back to start.
```

The result of this is that, though all increases of *stime* cause a small turn, larger continuous turns require the next increase in *stime* to come before the effects of the previous one have elapsed. In other words there is temporal summation of *stime* increases, constituting high pass filtering for rate of onsets. For auditory signals this sets the slowest rate at which successive syllables must occur for effective turns. The obstacle detection processing outlined above occurs quickly enough to fall within this time window and thus will cause continuous turning.

The original intention was for this temporal summation to be inherent in the motor properties of the vehicle, i.e., an *stime* increase would be translated directly into a turn, but the inertia of the vehicle would require more than one such turn to occur successively for an effective course alteration. However, the fastest rate for *stime* increase (and hence the minimum size of that increase for it to last till the next for summation) is determined by the rate at which RESPONSE can reliably detect onsets, which, as discussed above, is determined by the rate and accuracy of the auditory processing. This minimum time was far larger than the time required for an effective turn; furthermore there were minimal inertial effects in changing from moving straight ahead to turning around one wheel. Another option was to have movement unchanged (i.e. continuing forward) until the second increase within the period of the first occured, but again the necessary length of *stime* made this difficult as the robot could meanwhile move a fair distance.

Introducing a pause in response to the first *stime* increase solved most of the previous problems: the robot wasn't turning too far nor moving too far ahead before the next onset. It also had the advantage that the chance of correctly hearing that onset was improved by the robot being stationary. Having just a pause, without the small turn, would have been a more thorough implementation of the high-pass filter, but the interference levels in the acoustic environment were such that successive increases were often obscured anyway, and so some appropriate movement was necessary for every detected onset.

Note that the relative rates of syllable detection and turning are substantially different for the robot and the cricket—the robot with syllables of 300ms compared to the cricket's 15-25ms; the robot turning at around $90°/s$ and the cricket at an

average of $15°/s$. It would be wrong to conclude from the above discussion that the cricket generates an explicit move for every syllable: it seems more likely that if the high-pass filtering is based on temporal summation in the motor system, the input required to generate a turn occurs over a number of syllables. This is further discussed later.

A potential problem with independent modules controlling the motors is the possibility that they may interfere with one another in undesirable ways. In fact, the algorithm in RESPONSE makes it impossible for auditory input to demand simultaneous turns left and right, or for the obstacle detecting input to do so, but the two sensory sub-systems might compete. For example, if the left motor had been stopped due to a sound input requiring a left turn, but a left obstacle then led to a right turn being attempted, stopping the right motor, the robot would end up stopped completely. To avoid having an explicit representation of what had caused the turn, instead mutual inhibition between the LEFT and RIGHT modules was introduced, so that continuous turning in one would prevent response of other. In practice this gave a hierarchical advantage to responding to obstacles, because the rate of increase of the *stime* variable was faster and more reliable than the auditory response.

## 5.7 Measurement

To assess the behavioural abilities of the robot and compare them to the cricket, the arena paradigm was adopted because it can yield detailed information while maintaining fairly natural conditions. That is, the relationship of action to environment is not altered: making a move results in a change in the sound field as it does in a non-restricted situation. At the same time, it is possible to track the movements of the robot and to separate response to sound from avoidance of obstacles (both of which would be difficult in the more cluttered environment of the normal floor space) thus allowing a more thorough analysis of the behaviour to be made.

### 5.7.1 Apparatus

The floor surface on which the robot moves is linoleum. The arena measures 4x3.3 metres, and is constructed by blocking off this area of floor in the centre of the laboratory using 25cm high wooden planks. The laboratory itself is about 9x8 metres. The maximum arena area is dictated by the visual field of a video camera (focal length 4mm) suspended from the ceiling at a height of 3.7 metres. This arena is about 200x the size of the robot, which is rather small compared to cricket experiments where the arena is usually more than 1000x as large: for some of the experiments described in the next chapter, this limited the ability to distinguish useful results. A picture based on the view from the over-head camera (Figure 5–11) shows the arena, the loudspeaker that broadcast the sound, and the

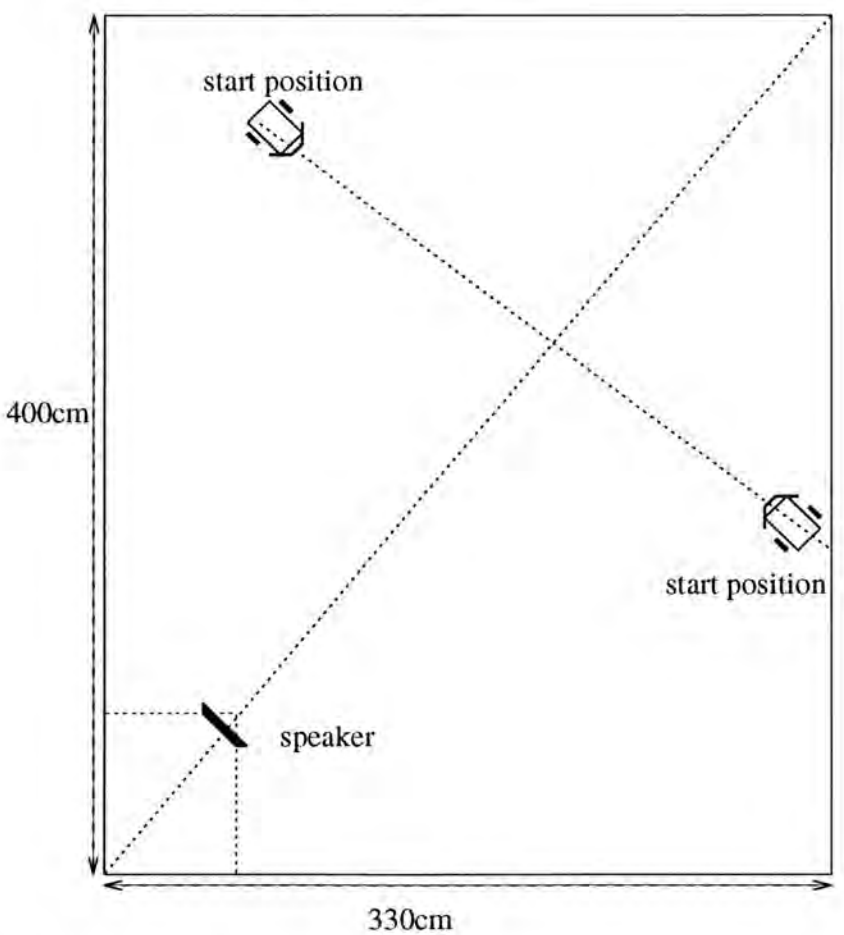



**Figure 5–11:** The arena as seen from the over head camera

two start positions used in the initial experiments (see next chapter).

The sound source is an 7-watt speaker (Sony SS-X1A) driven by a 2kHz oscillator circuit based on an 8038 function generator. The output is gated by the signal from a programable 8-bit microprocessor into temporal patterns of various kinds. The speaker response is not ideal and a certain amount of peripheral noise is associated with gating (visible as peaks in the auditory circuit response to patterned sound in Figure 4–6). Attempts were made to reduce this noise, but since it didn't prevent the robot from successful taxis these were not pursued. The speaker is usually positioned in a corner, sufficiently distant from the walls of the arena so that the robot can go behind it, to limit cases where the robot would reach the speaker as a consequence of wall following rather than auditory cues.

Both the walls of the arena and the walls of the lab are fairly good reflectors of sound, so it is fair to assume that the sound field only roughly resembles the ideal. In the cricket's natural environment the distortion is due less to reflection and more to absorbtion and atmospheric effects, which may place somewhat different requirements on the robustness of the phonotaxis mechanism. In the absence of the ability to accurately describe these respective sound fields, however, such effects cannot be predicted.

### 5.7.2 Dependent Variables

Four bright LEDs are attached to the top of the robot, and can be switched on and off by the internal program (as discussed above, this was used to signal whether the robot was moving forward or turning). With dim lighting, the bright point can be tracked, using a transputer based tracking program that produces image coordinates, estimated at pixel accuracy, at video frame rate (25Hz). Prior calibration was used to obtain a third order polynomial transformation to convert video coordinates to floor position coordinates, giving a positional accuracy of about $\pm$1cm.

Additional data is stored internally by the robot during a run (discussed above). The robot would stop when the memory for this data was full (when the array index reached 2000) which gave a little over 60 seconds running time as the maximum

trial length—in most instances the trial was ended by the robot reaching the sound source before this time had elapsed.

### 5.7.3 Procedure

Details of the procedures for specific tests are given in the next chapter, but the general procedure adopted was to start the robot from a specified position in the arena and allow it to run until it either hit the speaker or ran out of memory, recording the track on video and subsequently downloading the internal data, then starting the next run. Starting positions were standardised so that runs under different conditions could be compared. For most groups of comparisons, the different conditions were run alternately, rather than all of one then all of another, so that battery effects and other background variables, such as additional noise in the laboratory, would not bias the results. Occasionally runs had to be abandoned due to technical difficulties such as losing the program or a malfunctioning wheel or sensor, or because of failures in data recording; but otherwise all runs were included in the analysis. In other words, neither unexpectedly bad runs under usually successful conditions nor fortuitously good runs under usually bad conditions were excluded as exceptions.

### 5.7.4 Analysis

There are a number of possible ways of presenting and interpreting the data produced by the above procedures. As a particular aim was to compare results in some detail with cricket behaviour, the methods here were chosen so as to facilitate comparison, although the wide variety of methods used by neuroethologists could not all be included. An attempt was made to employ appropriate statistical tests where possible, but in many cases descriptive statistics are sufficiently informative about the behaviour.

The floor path was converted to polar co-ordinates centred on the speaker, as movement relative to the speaker is the behaviour of interest. Some of the basic

measures used below involve simple inspection of the path to determine how often, and how directly, the robot approached the speaker.

More complex processing was required to recover information about the heading of the robot relative to the speaker direction ($\theta$ in Figure 5–12). Note that this

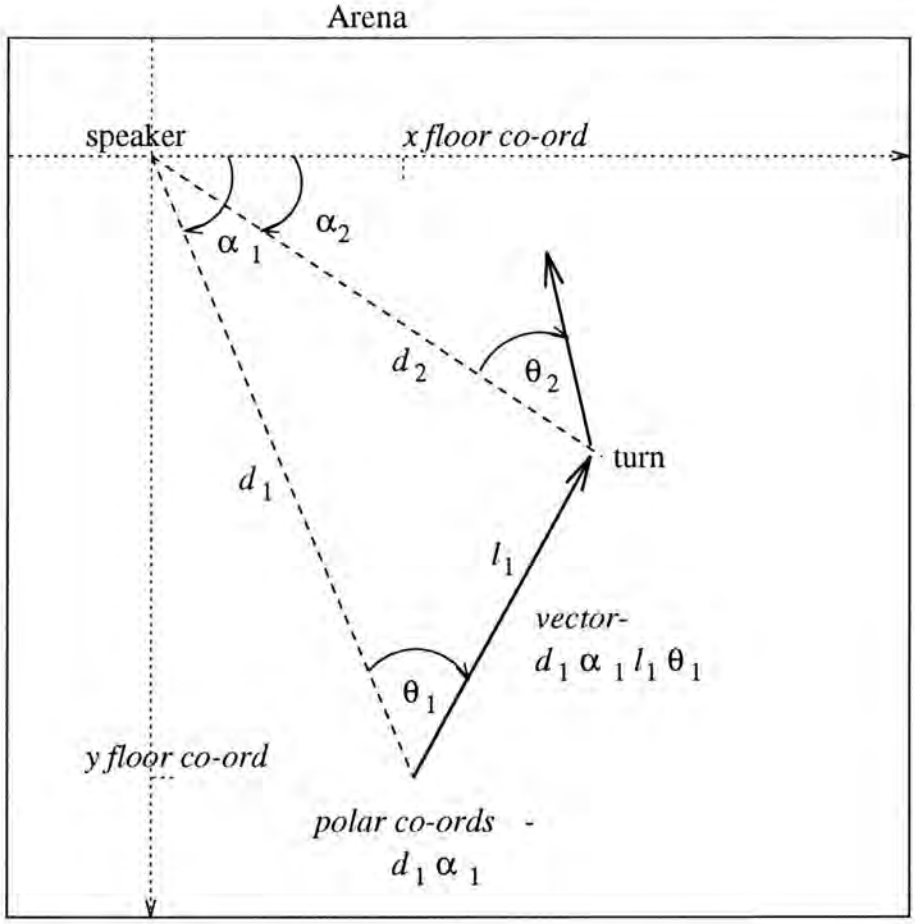



**Figure 5–12:** The track of the robot in the arena: amplitude and phase at each ear are determined principally by the distance $d$ and heading relative to the sound direction $\theta$.

heading is generally more important than the polar angle co-ordinate, $\alpha$, because with roughly spherical propagation of sound, the amplitude and phase at the two ears is independent of the latter angle, determined by the distance $d$ and heading $\theta$ alone. One way to calculate this heading is to find the direction between successive track points. As the points recorded were rather close, using successive points would lead to fairly inaccurate heading estimates, so in some cases the heading has been calculated after averaging over ten successive positions.

An alternative is to make use of the fact that the robot is known to move in (roughly) straight lines between turns. Then if the turn positions can be identified, a summary of the track consisting of the vectors between turns can be created. These vectors have four components: their origin (the turn position) in polar co-ordinates, and the relative direction $\theta$ and length of the move $l$ to the next turn position. The procedure for creating the vector summary for a track is illustrated in Figure 5–13. The time between points for the video track was accumulated to give a trial time that could be compared to the time data recorded internally. The start and end of trials could then be used to match up the internal and external data, and the internal record of when the vehicle turned could be used to extract the appropriate points in the video record (Figure 5–13iii). This process was done interactively so that the best match between the two data sets could be achieved. The averaged location of the points corresponding to a turn was calculated, and this series of positions used to create the vector set. The vector path is shown in Figure 5–13(iv), and it can be seen that very little information has been lost in this transformation.

The vector track can be used to compare paths in several ways. Plotting the set of vectors with all the origins at (0,0) gives a clear visual indication of the range of headings taken by the robot (v), and summing these to show the total distance travelled at each heading produces polar orientation diagrams directly comparable to cricket results. The vector mean (vi) for each track can also be calculated by finding the mean resultant of the vectors (each with length $l$ direction $\theta$):

$$\begin{array}{llll} \textit{mean resultant length} & \overline{R} = \sqrt{\overline{C}^2 + \overline{S}^2} & \textit{where} & \overline{C} = \frac{1}{n}\sum_{i=1}^{n} l_i cos(\theta) \\ \textit{mean direction} & \overline{X} = arctan(\overline{S}/\overline{C}) & & \overline{S} = \frac{1}{n}\sum_{i=1}^{n} l_i sin(\theta) \end{array}$$

The direction $\overline{X}$ indicates the general heading, and the length $\overline{R}$ represents how closely the vectors are clustered about that heading ($1-\overline{R}$ is a measure of variance, Mardia (1972)). A direct path to the speaker would thus have a vector mean of 0 degrees and length 1. Ordering of the vector means relative to this 'perfect' path has been used in several experiments to obtain non-parametric probability values for the difference between conditions.

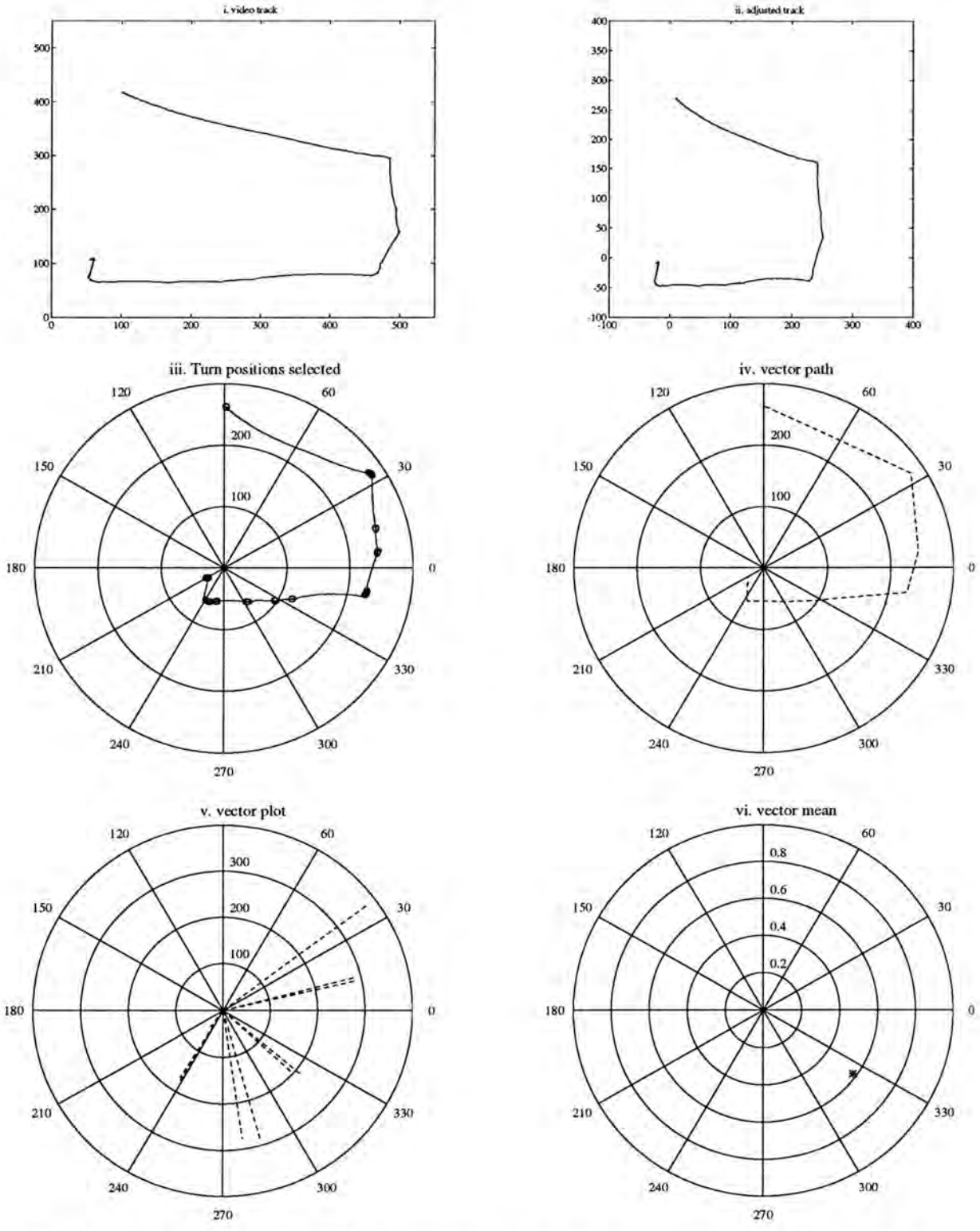



**Figure 5–13:** Processing of tracks: transformation from (i) video, to (ii) floor, to (iii) polar co-ordinates; identifying turns and (iv) constructing track of vectors between turns; (v) grouping the vectors and (vi) representing their direction and spread by a vector mean.

# Chapter 6

# Robot Phonotaxis: Results

## 6.1 Planned experiments

This chapter reports the results of a number of different tests and measurements of the behaviour of the robot described in the previous chapter. They were designed to examine the similarity (or otherwise) of the phonotaxis displayed by the robot to that of the cricket, to allow a reasonable evaluation of the viability of the mechanism as an explanation for cricket behaviour. They also demonstrate the range of capability of the mechanism as a means of getting sound-locating behaviour in a robot.

The experiments carried out were:

- Testing the ability of the robot to move towards the ideal sound, compared with behaviour with no sound. The tracks taken were analyzed for directness, correctiveness, zig-zag characteristics and intensity dependence—these are measures used to characterise cricket behaviour.
- Examining the behaviour of the robot with different syllable rates for 'recognition' effects: *i.e.* for failure to perform taxis if the sound was not ideal.
- Looking for possible effects of chirp patterns on the behaviour.
- Testing whether the sound could still be effectively located when the environment was made more complex by adding obstacles to prevent a direct route. Also testing behaviour over a larger area (the entire lab rather than just in the arena).
- Looking at the results with two sound sources to see if this 'simple' mechanism could reproduce some of the more 'complex' cricket results.
- Testing if the robot could perform taxis with 'one ear'.

## 6.2 Taxis

The most obvious first requirement is to show that the robot equipped with this sensory-motor mechanism can perform phonotaxis, that is, does move towards a sound source. The basic paradigm used in this and the next section was to run the robot alternately from two starting points on opposite sides of the arena, each about 2.5 metres distant from the speaker, with the robot facing towards the other side, at right angles to the speaker direction. Thus the robot would have to alter its heading to go towards the speaker (if the robot is started facing the speaker, then a program consisting of the command 'go forward' will be the most efficient way to reach the speaker, which would make it difficult to rate the relative success of the phonotaxis program). The trials described in this and the next section were run in a mixed order (as mentioned above) to minimise bias.

In this section I will compare tracks performed without sound (4 trials) to those with a sound source producing the 'ideal' syllable rate of 1.6 Hz as defined above (10 trials). The simplest measure of phonotaxis is to see whether the robot does reach the speaker before the trial is ended by a full memory (that is, within about 60 seconds). Without sound one trial failed to reach the speaker, one had the memory run out when just behind the speaker, and two of the trials ended at the speaker after detours around the arena (Figure 6–1). In all ten trials with sound the robot reached the speaker (Figure 6–2).

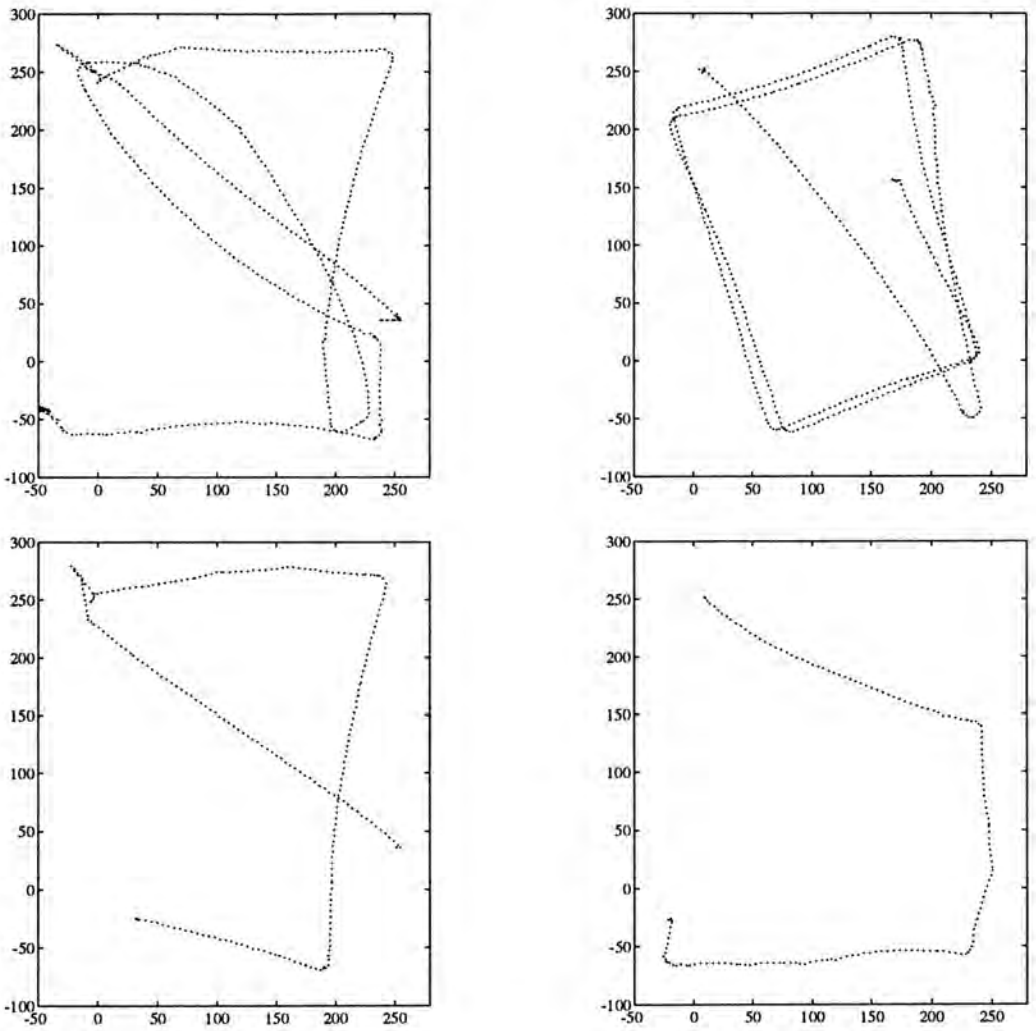



**Figure 6–1:** Without sound: the robot is started from the left or right, and moves around the arena, turning away from walls. The run is stopped if it reaches the speaker location (0,0) or if memory is full (about 60 seconds).

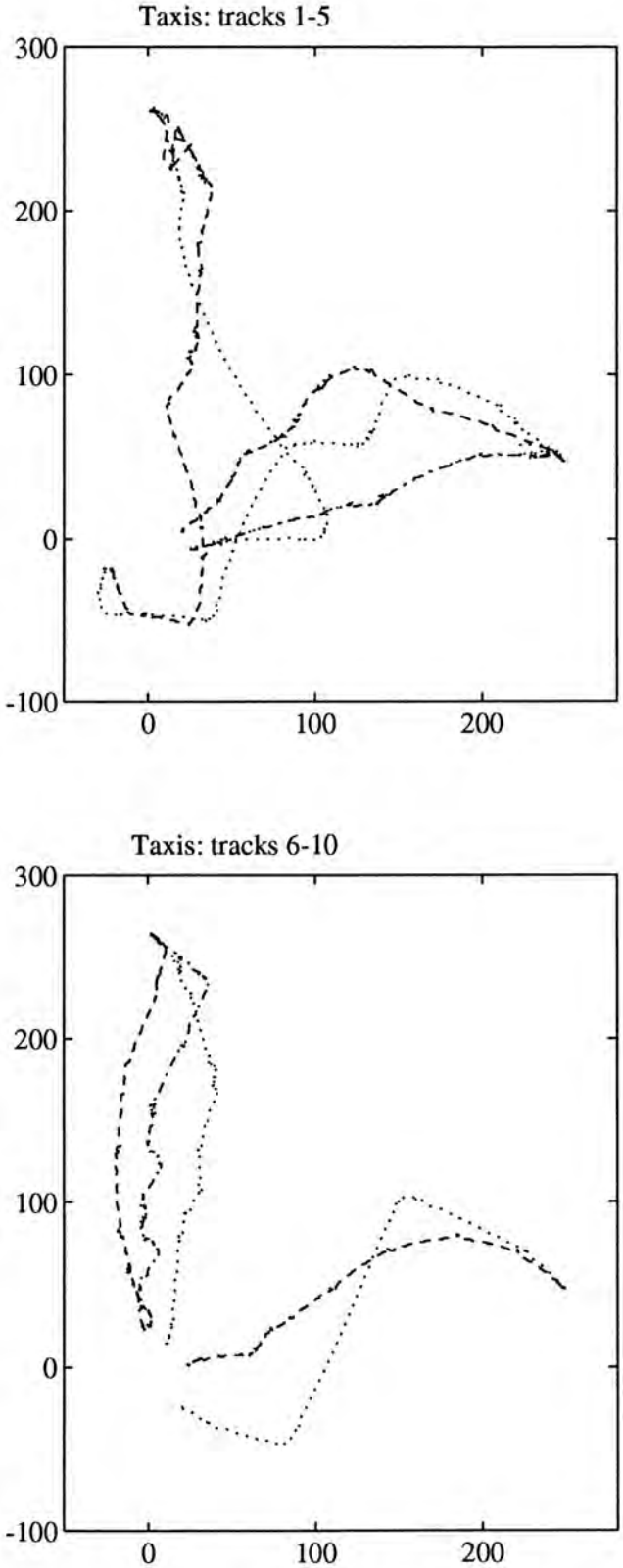



**Figure 6–2:** Taxis to 'ideal' signal: 1.6kHz syllables of 2kHz tone. Ten trials start alternately from two locations. Trial is stopped when robot hits the speaker (located at 0,0).

The difference in the paths is quite apparent in these pictures, but a more thorough evaluation of the behavioural difference can be made by using measures for characterising successful taxis that have been adopted in cricket experiments. Using the arena paradigm, Stout and McGhee (1988) compare track lengths and classify as 'direct' a path that is less than twice as long as the straight line distance from the start point to the speaker; and also look at the angle of preference in a polar orientation diagram. For the treadmill, Schmitz *et al* (1982) note the average direction indicated by mean vectors as well as looking the heading angle against time, the dependence of angular velocity (turning rate) on prior heading, and the improvement in accuracy with the amplitude of the sound. Thorson *et al* give"criteria for clear tracking" consisting of four features: i) corrective meandering centred on the speaker direction; ii) characteristic stop-start walking; iii) no more than $\pm 60°$ deviation from the speaker direction; iv) following the sound when it is switched between speakers.

### 6.2.1 Direct course, within $\pm 60°$

Again, by examination of the tracks shown in Figure 6–2 it is clear that the course taken to the speaker is usually fairly direct and involves little wandering, and that mostly near the speaker. The ratio of track lengths to direct distance ranges from 1.017 to 1.806 *i.e.*, all are 'direct' paths. For no sound, ratios range from 3.307 to 10.302. Comparison of the sound tracks with those produced by crickets (Figure 6–3) shows some similarity, although it should be kept in mind that crickets vary widely in their tracking behaviour, both within individuals and between individuals, and under different experimental paradigms.

The vector plot (see 5.7.4 for an explanation of how these are produced) of all the tracks also shows clearly the directionality of the tracks to sound compared to those with no sound (Figure 6–4, left); the vectors are centred on 0° and are almost entirely contained within $\pm 60°$. Summing the length of the vectors in each 10° range produces a 'polar orientation diagram' as used by Stout *et al* (1983). The cricket plots (Figure 6–3) are more strongly directional than for the robot

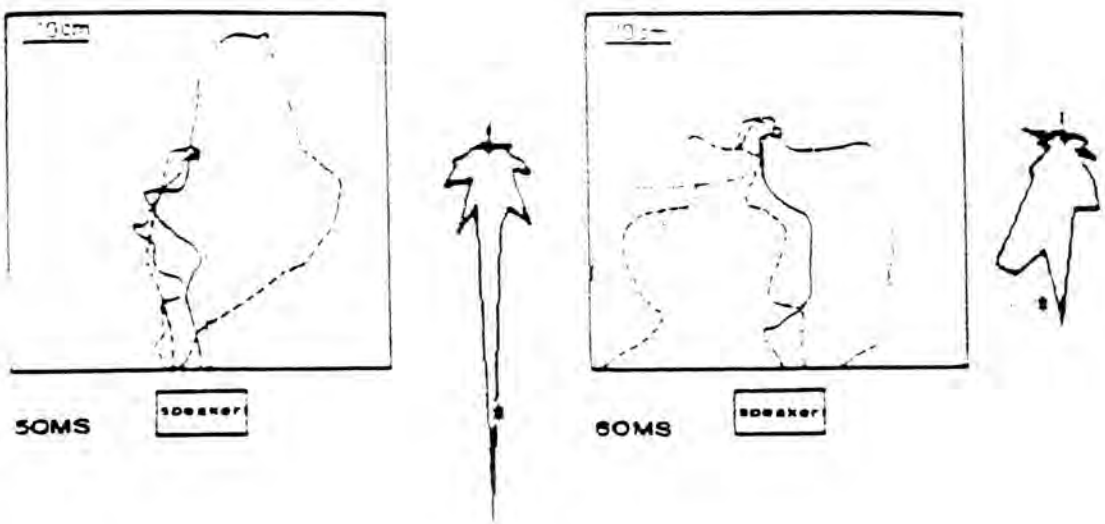



**Figure 6–3:** Cricket tracks in arena and polar orientation diagrams at close to ideal syllable periods (50ms and 60ms). From Stout *et al* (1983).

(Figure 6–4, right), but the robot plots do clearly show taxis towards sound as opposed to less directed movement when sound is absent.

### 6.2.2 Corrective meandering

Corrective meandering refers to the zig-zag course that crickets perform to the sound, that is, the angle to the sound varies about the midline but turns tend to take the cricket back towards it. This is one aspect of the behaviour that tends to suggest phonotaxis is largely reactive.

One way to look for this characteristic in the robot path is to plot against time the angle relative to the speaker that the robot is heading (plots of this kind are immediately provided by the recording methods used in treadmill experiments). For this I have used tracks produced when the robot started in the opposite corner to the speaker (baseline behaviour for the obstacle experiments described in 6.3), rather than the ten trials discussed so far, to avoid the distortion caused by the robot *starting* by facing the wrong direction, and because a greater distance was

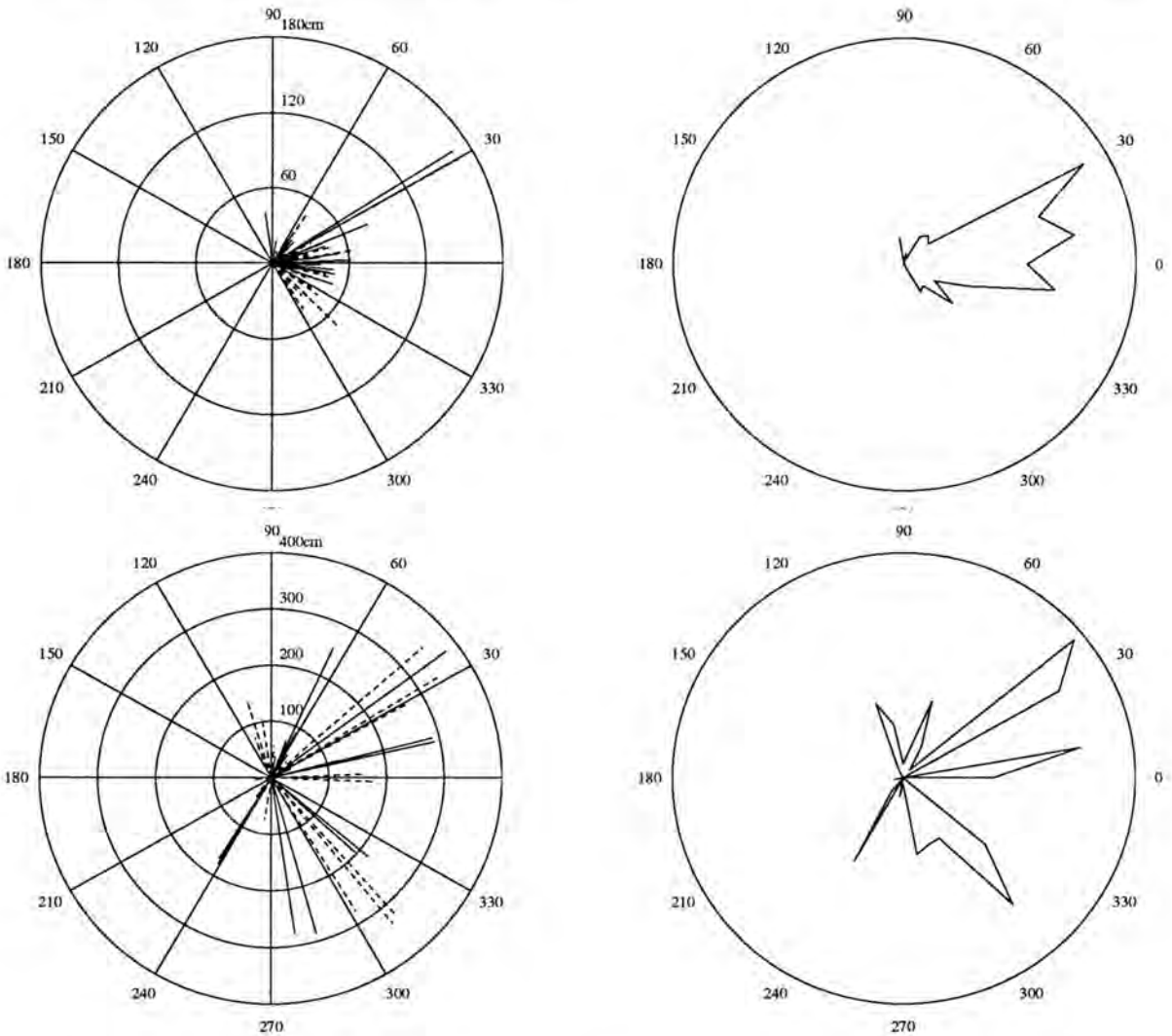



**Figure 6–4: Left:** the set of directional vectors taken from the tracks in all trials. **Right:** total distance moved in each 10 degree interval (polar orientation diagram). **Top,** ideal sound, **bottom,** without sound.

covered. In Figure 6–5 the position of the robot has been averaged over each 10 measurements (about 400ms) and the angle between each resulting position estimate calculated. The resulting trace does show zig-zag characteristics similar to continuous cricket walking on the treadmill.

Another way to evaluate the 'correctiveness' of turns is to look at the change in angle or angular velocity relative to the previous walking direction. Schmitz *et al* (1982) call this the 'characteristic curve' and present some evidence for a sinusoidal dependence of angular velocity on heading direction. In Figure 6–6 the difference

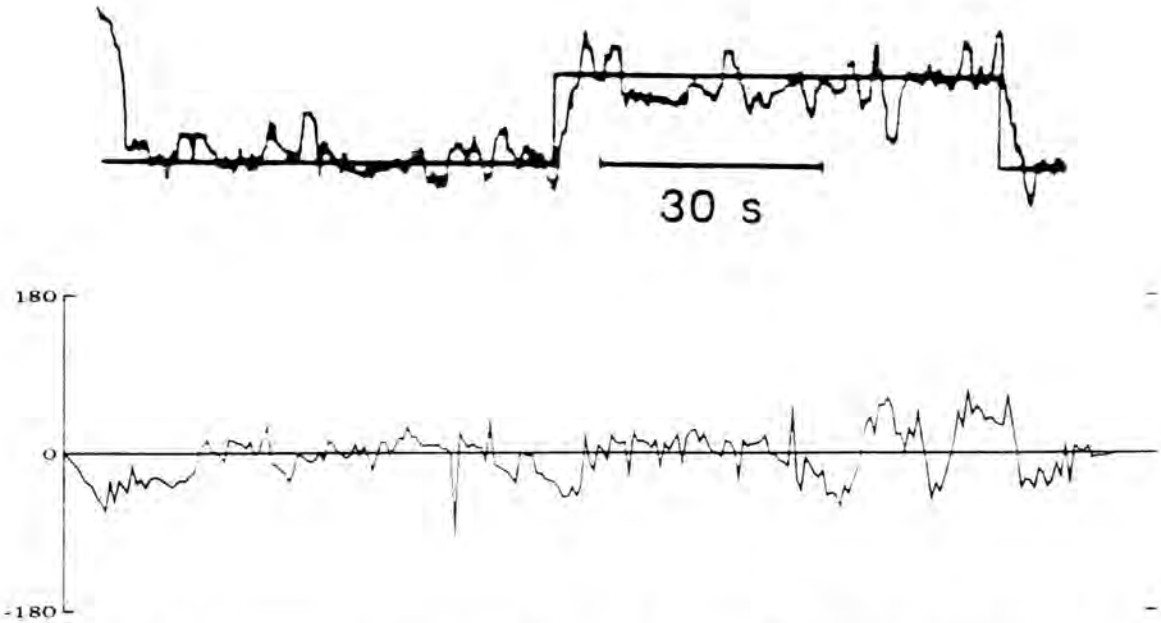



**Figure 6–5:** Corrective meandering. **Top:** cricket on treadmill (Thorson *et al*, 1982), includes switch in speaker direction (solid line). **Bottom:** robot in arena, speaker direction 0°.

in angle between successive vectors (the size and direction of the turn) has been calculated for each direction (grouped in 20° intervals) headed immediately before the turn. The mean and standard deviation of these turns is shown and can be compared with the equivalent angular velocity plots from Schmitz *et al* (1982). There is some similarity: turns for positive directions are negative, and for negative directions are positive; the size increases towards 60° then levels off (although, as already discussed, there are few measurements outside ±60°). Between −60° and +60° degrees, correlation between the variables is −0.4048.

### 6.2.3 Follows switch

The fact that the cricket follows a switch in speaker direction can be taken to indicate two things: first that the directed walking is towards the sound rather than coincidently in the right direction, and second that it is basically reactive in nature, the cricket responding to the current sensory situation without apparent reference to what preceded it. For the latter aspect, the program in the robot is

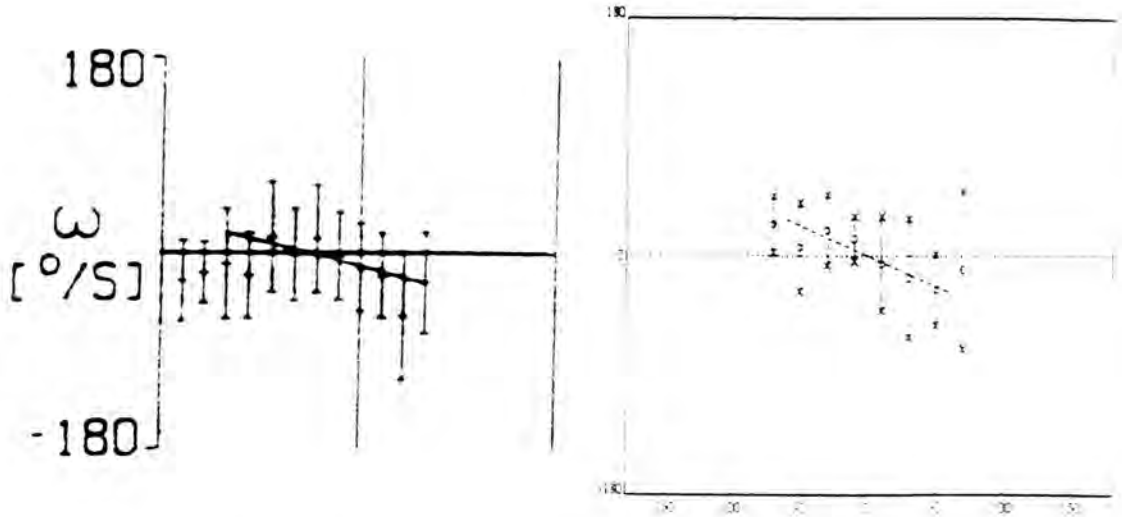



**Figure 6–6:** Characteristic curve (mean and variance of angular velocity for different heading directions). **Left:** cricket on treadmill with medium intensity sound (Schmitz *et al*, 1982); regression slope is −0.254. **Right:** robot in arena, averaged over ten tracks: regression slope is −0.4586.

designed to be reactive so this need hardly be shown. For the former, the fact that the robot has been started from different places, facing different directions, but still goes to the speaker makes it clear that the path does depend on the direction of the sound relative to the robot.

In later experiments with two speakers (section 6.5) some attempt was made to show that switching between the speakers would cause the robot to change direction. Two paths showing this are plotted in Figure 6–7. However the experimental set up made it difficult to record precisely the point in the track where the speaker change occured so as to measure how long it took for the robot to change direction. Also, within an arena the robot will be going between speakers at 180° separation *i.e.* the new sound will be behind it. A sound from behind has the same effect as one directly in front (although travelling away from the sound is an unstable situation) so the recovery time will be affected: in tests with the

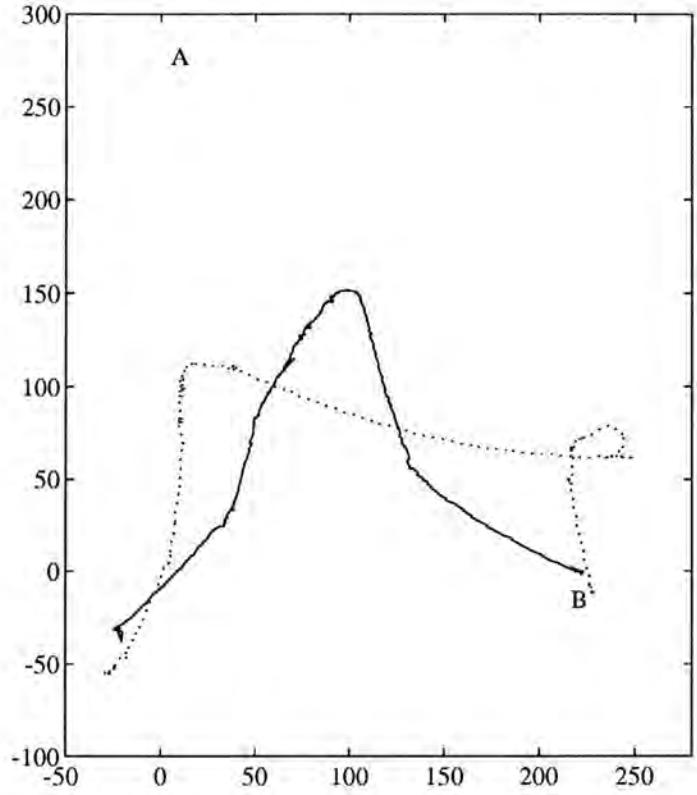



**Figure 6–7:** Tracks with robot starting in lower corner $(-50, -50)$ and sound switching from speaker A to speaker B when robot nears centre of arena

robot this was sufficient to have the robot reach a wall and this interfered with clear tracks in response to sound.

### 6.2.4 Moves in a characteristic manner

In several papers the 'characteristic' nature of cricket's movments during phonotaxis is noted. On the treadmill, phonotaxis tends to have quick movements and frequent stops, whereas without sound walking is slower and more continuous (Weber *et al*, 1981). The onset of this characteristic movement has been used to indicate 'recognition' in the absence of a direction of sound (with the sound broadcast above the animal). In the arena, the behaviour of crickets not doing phonotaxis is usually rapidly to approach the nearest wall and try to climb it, or follow it round the edge (own observations, confirmed by tracks in Stout and McGhee (1988)).

The robot's movements with the sound are quite noticeably different to those without sound, although this is not all that surprising given the program. In unrecorded pilot trials with different speaker positions, this difference was also noticable even if the robot was not successfully approaching the speaker: *i.e.* in the absence of consistent directional cues it would still perform pauses and turns due to inconsistent responses. This is noted because I have previously suggested such inconsistent responses may explain undirected but 'characteristic' behaviour in the cricket, rather than this behaviour being evidence of an independent recognition mechanism.

### 6.2.5 Intensity dependence

As mentioned in chapter 4, Schmitz *et al* (1982) have argued that the intensity dependence of phonotaxis (with accuracy improving for louder sound) is evidence against a latency mechanism. I have already suggested there that this argument is flawed, but it seems worthwhile to attempt to show that the latency mechanism employed here *can* produce an intensity-dependent difference in behaviour. As the robot approaches the speaker in the arena, intensity of sound depends on distance, rather than being an independent variable as it is for the treadmill paradigm. However, given the reflective environment the intensity is probably only loosely correlated with the distance for many positions in the arena.

Looking at the distribution of direction before turning against the distance from the speaker (Figure 6–8), there does appear to be less deviation from the zero direction as the distance decreases from 300cm to 70cm (there are some outliers at 100cm that possibly correspond to avoidance of walls), and then an increase which is perhaps explained by the fact that near the speaker there are obstacles which interfere with the movement (this is also approaching the near sound field for the speaker). This effect is not very strong, but does give some indication that accuracy may improve with intensity for this mechanism. In later trials with the robot free to move round the entire lab (see 6.4) it certainly appeared to perform worse when at a further distance from the speaker.

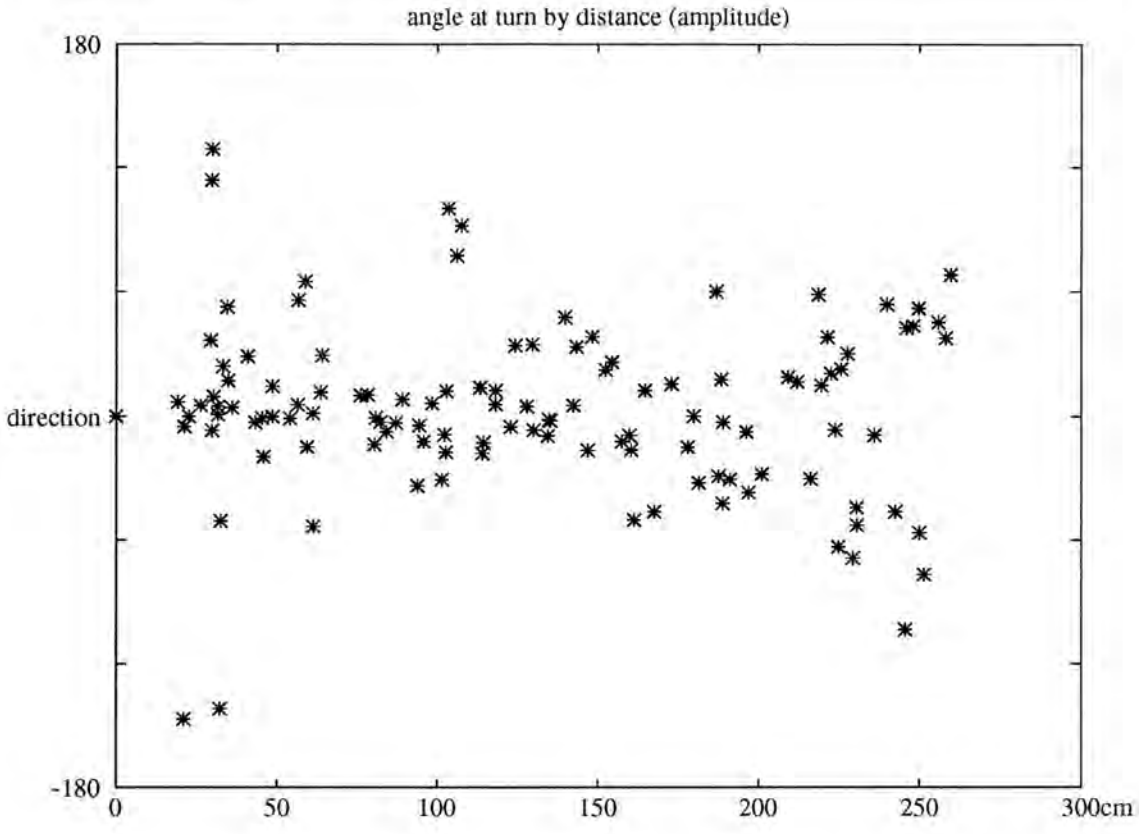



**Figure 6–8:** Direction before turning by distance from speaker

### 6.2.6 Discussion

Previously I have indicated that taxis is in general a fairly robust mechanism and not all that difficult to implement, so the fact that the robot can find the speaker in a fairly simple environment is not in itself strong evidence that the mechanism is the same as that in the cricket: there are several possible mechanisms that may also achieve this result. It could be argued that using the cricket as inspiration did lead more quickly to a working solution to the problem of approaching sound sources in particular—the use of phase cancellation and reactive rather than planned movements are two useful techniques thus derived.

But a more thorough analysis of the robot's behaviour has shown a number of characteristics that quite closely resemble the behaviour of the cricket. Few of these were explicitly programmed: the zig-zag course, the restriction to within $\pm 60°$, and the characteristic curve emerged from the interaction of the mechanism with the environmental situation. This close similarity is more surprising when it

is considered that the robot and cricket have substantially different motor modes; and is a stronger indication that the mechanism might in fact be a reasonable model of the cricket's neural system. However it would be necessary to examine whether or not other mechanisms for taxis also produce similar results before drawing firm conclusions.

## 6.3 Recognition

One of the main influences on the design of the mechanism was that the ability to find the sound should inherently depend upon that sound having a particular frequency and temporal pattern, thus creating the impression that the robot 'recognises' the correct sound for phonotaxis. To show that such an effect does arise from the mechanism described it is necessary to examine the robot's behaviour to sounds that vary in specific ways from the ideal. Although the appropriate behaviour may be expected because it was designed that way, it is worth demonstrating that the mechanism *does* work in the real situation.

To demonstrate that at least some temporal patterning is required, the behaviour with continuous sound at 2kHz can be compared with the taxis to sound described above: from the tracks (Figure 6–9) it is clear that this behaviour differs little from the results with no sound.

### 6.3.1 Preferred syllable rate

As discussed in chapter 4, it is fairly well established that the most critical feature for phonotaxis in the cricket is the rate at which the syllables repeat. The phonotactic response decreases for syllable rates faster or slower than the ideal. To test this in the robot, a series of trials was performed under the paradigm described above (with two starting positions, trials ended by reaching speaker or running out of memory, tracks recorded and processed) using syllable rates of 5Hz, 2.5Hz, 1.6Hz, 1.25Hz, 1Hz and 0.8Hz (that is, with a 50/50 duty cycle, and syllables of

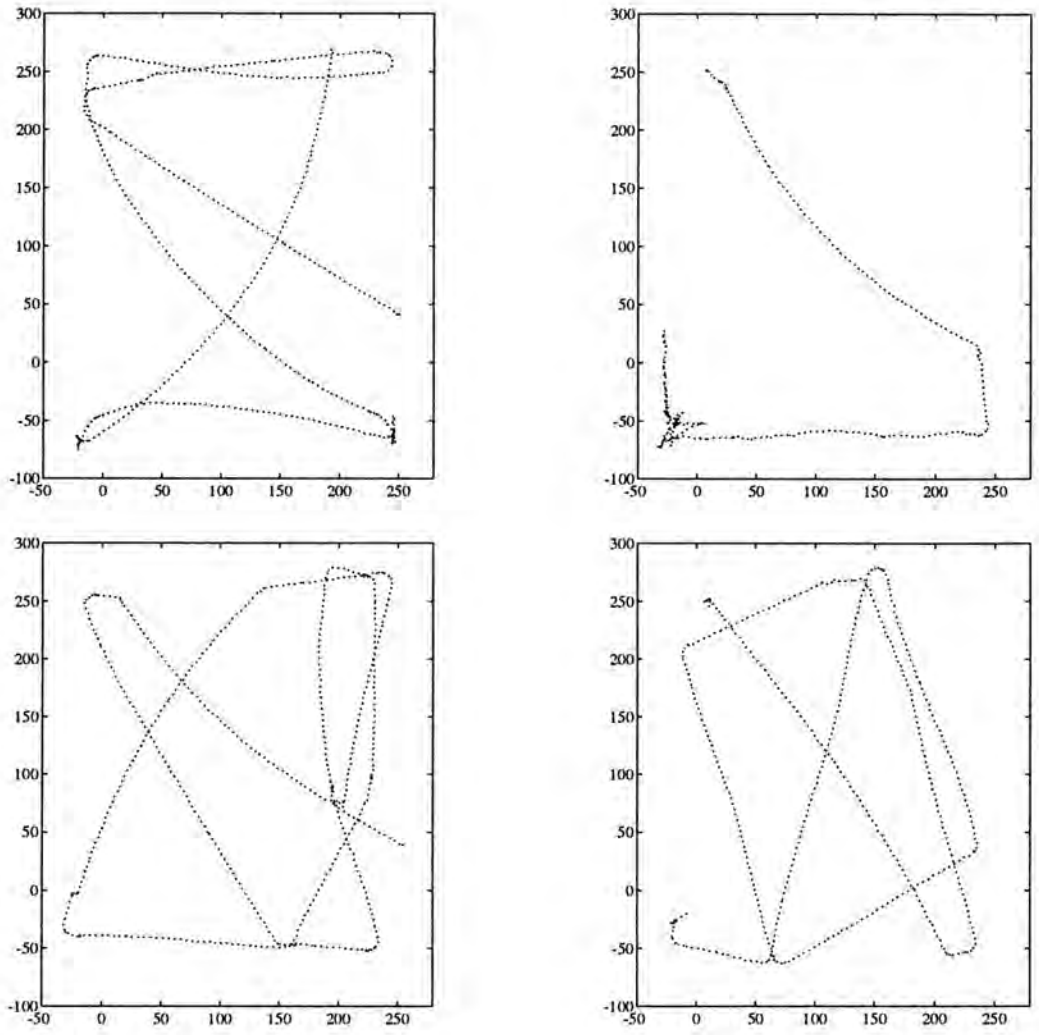



**Figure 6–9:** Four tracks with a continuous 2kHz tone, speaker at (0,0).

length 100, 200, 300, 400, 500 and 600 milliseconds). For each rate, ten trials were run, in pairs from the two starting positions: pairs of trials at different rates were mixed together. Figures 6–10& 6–11 shows the best, worst and median track for each syllable rate, and the vector plot for all the tracks at that rate (because of the complexity of processing to extract vector plots, these were not done for the fastest and slowest rates, as the other rates were sufficient to show the trend).

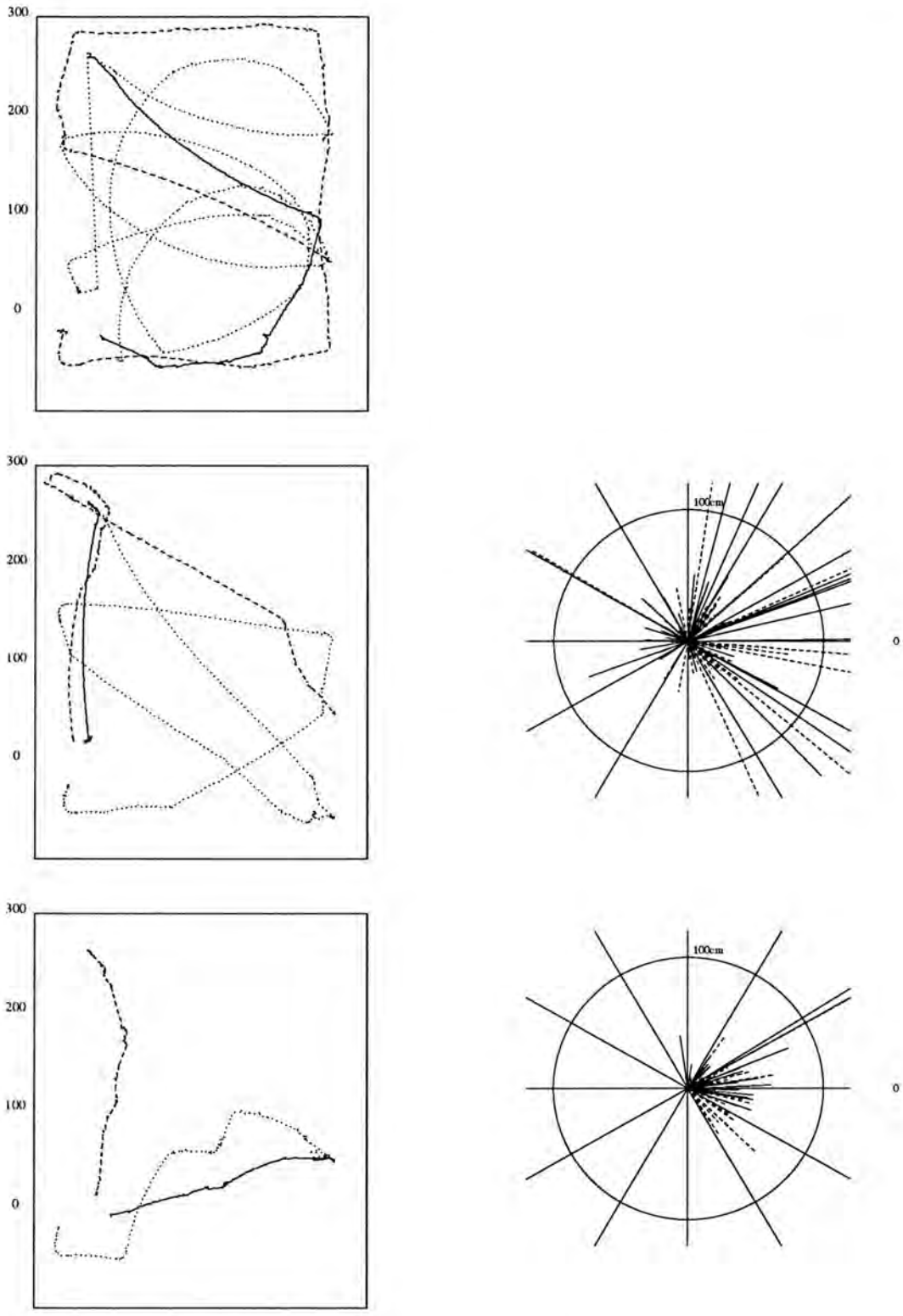



**Figure 6–10:** Tracks at 5Hz (top), 2.5Hz (middle) and 1.6Hz (bottom); shortest (solid line), median (dashed), longest (dotted). Collected vectors for ten trials at respective rates; starting from left (solid) and starting from right (dashed).

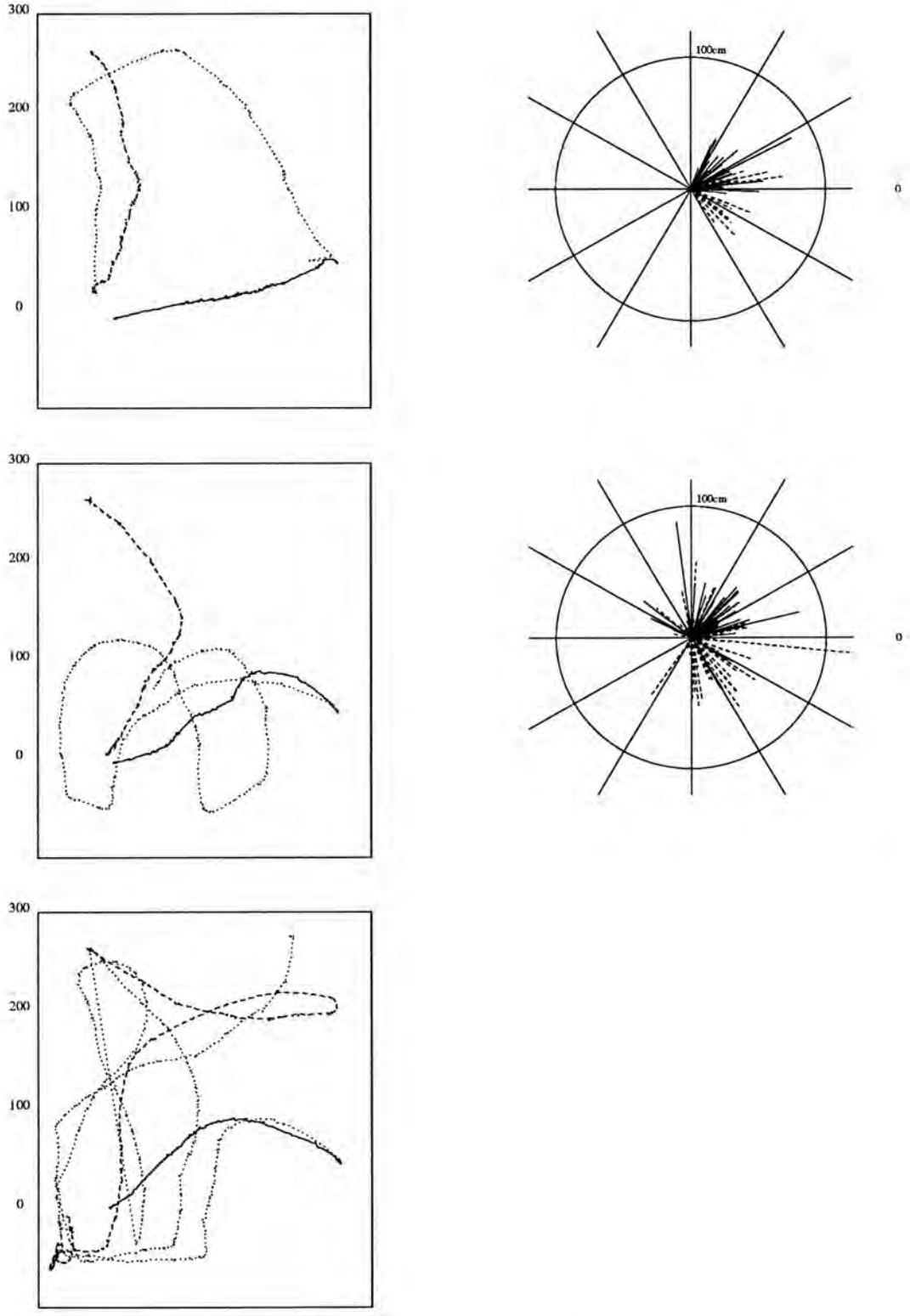



**Figure 6–11:** Tracks at 1.25Hz (top), 1Hz (middle) and 0.8Hz (bottom); shortest (solid line), median (dashed), longest (dotted). Collected vectors for ten trials at respective rates; starting from left (solid) and starting from right (dashed).

It is apparent that taxis is almost non-existent at 5Hz, and is inefficient at 2.5Hz—the best track here is the result of a single turn towards the speaker followed by a straight, uncorrected, run. At rates slower than the ideal (1.6Hz) the robot more reliably approaches the speaker but finds it through repeated small turns that make up a curving track: for 1.25Hz the curvature is sufficiently high to result in reasonable taxis, at 1Hz and 0.8Hz the decreasing size of the curvature starts to adversely affect the directness of the path to the speaker. This difference in the paths is evident in the shorter vectors, and also in the fact that the start position has more influence on the path as is evident in the different range of vectors from the left start (solid) and right start (dashed).

It is worth noting that Thorson *et al* (1982) explicitly exclude tracks of this latter kind, "segments of circling (without corrective meandering) which bring the animal near the direction of the active loud speaker" under their tracking criteria, which may partly account for the "greatly reduced tracking" they report outside the preferred syllable range. They do not present any details about what tracks looked like outside that range, but Stout *et al* (1983) note that "Both *A. domesticus* and *G. campestris* were more tolerant to increases in the syllable period above the conspecific modal values than to decreases in syllable period", where period is defined as time between onsets of syllables. Thus slower rates cause a less dramatic reduction of taxis than faster ones in crickets as well as in the robot.

To test whether tracking was really better at the ideal rate, the difference between tracks at 2.5Hz, 1.6Hz and 1Hz was examined statistically. As described in the methods, a vector mean can be calculated that represents the direction and directness of a track: the means for each of the ten tracks at these rates have been plotted in Figure 6–12. The ideal mean would be direction 0, length 1, so the distance of the means from this point was used to find a relative ordering from most direct to least direct. Using the Mann-Whitney *U*-test (Mendenhall *et al*, 1981) with $n_1 = n_2 = 10$, it appears that tracks at 1.6Hz are significantly more direct than at either 2.5Hz ($U = 17, p < 0.0057$) or 1Hz ($U = 11, p < 0.0010$).

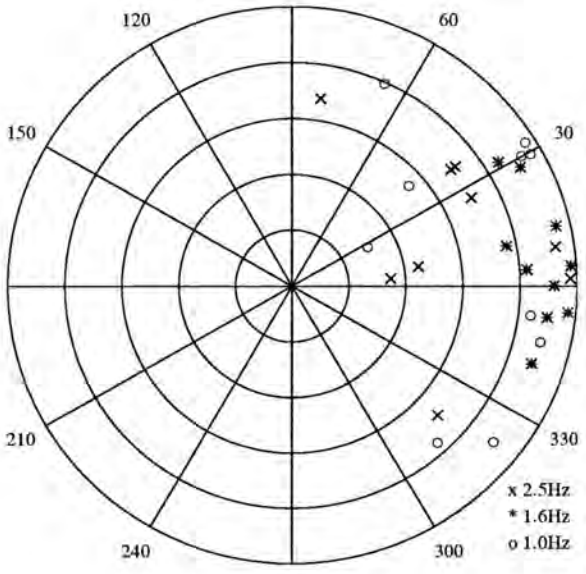



**Figure 6–12:** Mean vectors for each of the ten trials at three different syllable rates: compare distance of means from the ideal, direction 0, length 1.

### 6.3.2 Syllable rate vs syllable length

In the previous experiment the duty cycle was constant. Consequently, the songs used differed in syllable length as well as syllable rate. To rule out the possibility that syllable length is a critical factor in the significant results, two further sets of trials were run, using short syllables (200ms) at the 'ideal' rate of 1.6Hz, and using 'ideal' syllables (300ms) at the slower rate of 1Hz. Again ten trials of each were recorded, and the resulting mean vectors are plotted in Figure 6–13, along with the original 300ms/1.6Hz results. Following the same comparison procedure, there was no evident difference between 200ms and 300ms syllables at 1.6Hz ($U = 38, p > 0.2$), but a significant difference between 300ms syllables at 1.6Hz and 1Hz ($U = 19, p < .0093$). So the more direct tracks do appear to be associated with the rate, rather than the length, of the syllables.

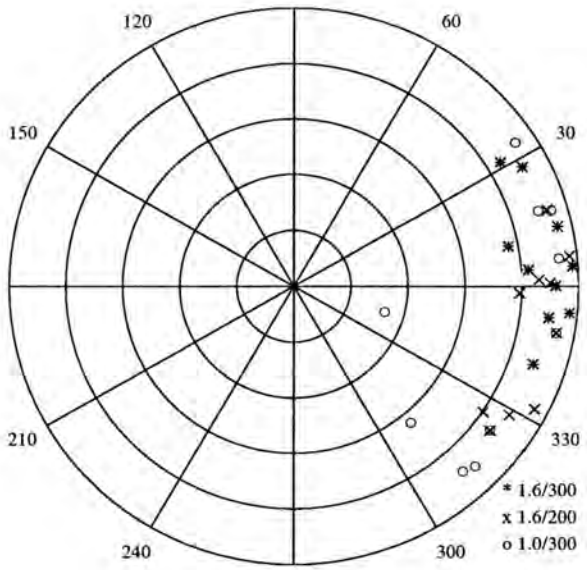



**Figure 6–13:** Mean vectors for each of the ten trials at different syllable rates/syllable lengths

### 6.3.3 Discussion

It is evident from the above experiments that a difference in behaviour to alterations in the temporal pattern of the signal can occur as the result of a latency-based comparison mechanism for taxis. Also, those differences resemble the effects attributed to 'recognition' in the cricket. This implies that the syllable-rate dependency of the cricket's response can plausibly be explained by a combination of slow auditory neural response (effective low pass filtering) and temporal summation in motor neuron response (effective high pass filtering), where both are inherent features of the onset-dependent latency-comparison mechanism. An independent 'recognition' function is not necessary to explain the results.

The implementation of this filtering in the robot is probably somewhat simpler than is likely in the cricket. In particular, the summation that leads to longer turns is probably too direct, because the speed of the robot relative to sound processing

made motor response to individual syllables necessary. The faster rate in the cricket would make it viable to require the integration of two or more consecutive syllable responses before a turn was made, *i.e.* response to individual syllables may be negligible. Also the control of a six-legged turn would require more complex neural signalling than 'stop one side': the form of neural signalling required to generate turns would affect the implementation of the high-pass filtering.

Furthermore there are some aspects of the cricket behaviour that the mechanism, as it stands, would not reproduce. In particular, Thorson *et al* (1982) report that a song with 90% duty cycle (i.e. long syllables) will still produce clear tracking in the cricket, whereas the robot fails to respond when gaps between syllables are small. This suggests that attributing the failure of taxis at fast rates to the time constant of the AN1 neuron needs further elaboration: why is a 7.5ms gap at 15ms syllable interval not detectable, yet a 5ms gap is sufficient for taxis with a 30ms interval? One possibility is that the 'time constant' is not constant: there is some evidence that over the length of a syllable the firing rate decreases (a slight adaptation effect?), so if the recovery time depends on the firing rate, it would be shorter after long syllables than shorter ones. An effect of this sort could be implemented in the robot but would require careful estimation of parameters to work properly: clearer evidence about the nature of the neural response in the cricket would be required.

Some of the tracks to 'non-ideal' rates actually resulted in faster approach (*i.e.* quicker, though less direct) to the speaker than at the ideal rate. Once the robot was headed in roughly the right direction, moving forward between small corrective turns could be quite efficient, whereas at the ideal rate there was a certain amount of unnecessary vacillation. This behaviour is interesting because it suggests a possible role for the chirp structure in the natural cricket song: within syllable groups the cricket can adjust direction and between them it can move forward, in roughly the right direction, thus increasing the speed of approach.

### 6.3.4 Chirps

This idea was tested using the same two-start paradigm but with three-syllable chirps followed by silence for the same length of time. However, the small size of the arena limited the results—the robot could move far enough between chirps to substantially alter the heading relative to the speaker. Nevertheless, as shown in Figure 6–14, these chirps could lead to effective taxis, with fewer turns and

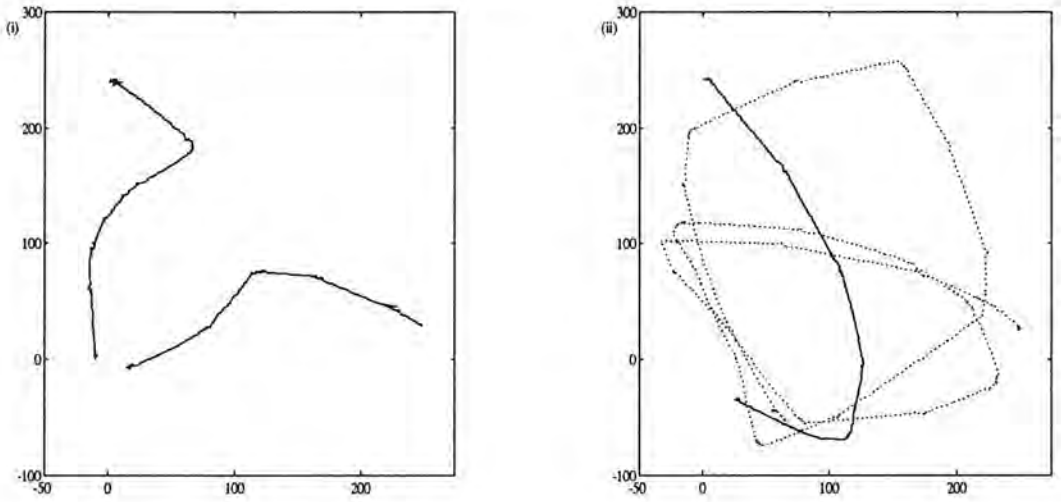



**Figure 6–14:** (i) Two tracks with three-syllable chirps.(ii) Two tracks with a chirp pattern only (no syllables).

consequently faster paths than the continuous syllable train. This effect would probably be more noticable over a larger distance, or with a slower robot.

Stout and McGhee (1988) have shown some crickets would approach a speaker reproducing only the chirp pattern of the song. This is equivalent to very long syllables, so it might be possible that, as for long syllables above, an animal starting in the right direction might reach the speaker despite only being able to make small adjustments to the course. Such an effect is shown here for the robot using same chirp pattern as before but without syllable patterning—effectively 1500ms syllables or a rate of 0.3Hz. Small corrections plus turning the right way from a wall result in reaching the speaker in one example, whereas in another the robot repeatedly overshoots the speaker and takes a long time to recover the right direction.

One interesting implication of this explanation for the chirp is that it predicts different results for different paradigms for measuring taxis. For a fixed cricket, where the attempt to turn in a sound field is measured, continuous syllables would be expected to produce a stronger (because continuous) result. On the other hand, the speed of the track in an arena would be faster with chirps than without. There does seem to be some evidence for this in crickets: with *Teleogryllus* showing a 'preference' for trills over natural song for tethered flight experiments, but not in arena experiments (Pollack and Hoy (1981)— they discuss potential reasons for the difference but don't mention this particular possibility); and the evidence for significance of chirps usually coming from arena experiments.

## 6.4 More complex environments

### 6.4.1 Obstacles

Given the criticism in earlier chapters of unrealistic tests of mechanisms making doubtful the validity of conclusions, it is necessary to test whether the success of the phonotaxis mechanism described is dependent on the fairly simple experimental situation of a small, empty arena. It is known that the cricket can do phonotaxis in a complex environment (although exactly how well is rather difficult to assess in field studies). The complexity of the robot's environment can be increased by adding obstacles between the start point and the speaker, which the robot must find its way around. Note that such obstacles will also increase the complexity of the sound field through distortion and reflection.

The obstacles used for the trials in Figure 6–15 were 25x25cm boxes of the same height and material as the walls of the arena. They were placed a minimum of two robot lengths (60cm) apart (so the robot might have some chance of going through the gap rather than treating it as a continuous wall—recall that it only has two binary sensors for obstacle avoidance). They were arranged so that the robot would need to make detours to reach the speaker when travelling from one corner to the other. In another set of trials (Figure 6–16), a 1.5metre long wall

made from the same kind of plank as the arena walls was placed so as to block the direct path from corner to corner.

As was discussed in the Chapter 5, it was found necessary to modify the motor control interaction to prevent interference of turns away from obstacles with turns towards sound. Apart from this no other alterations were made to the program.

The more detailed forms of analysis are less meaningful in this situation because turns are affected by obstacles as well as sound. However, it is evident from the tracks Figure 6–15, that the robot did successfully find its way to the speaker in

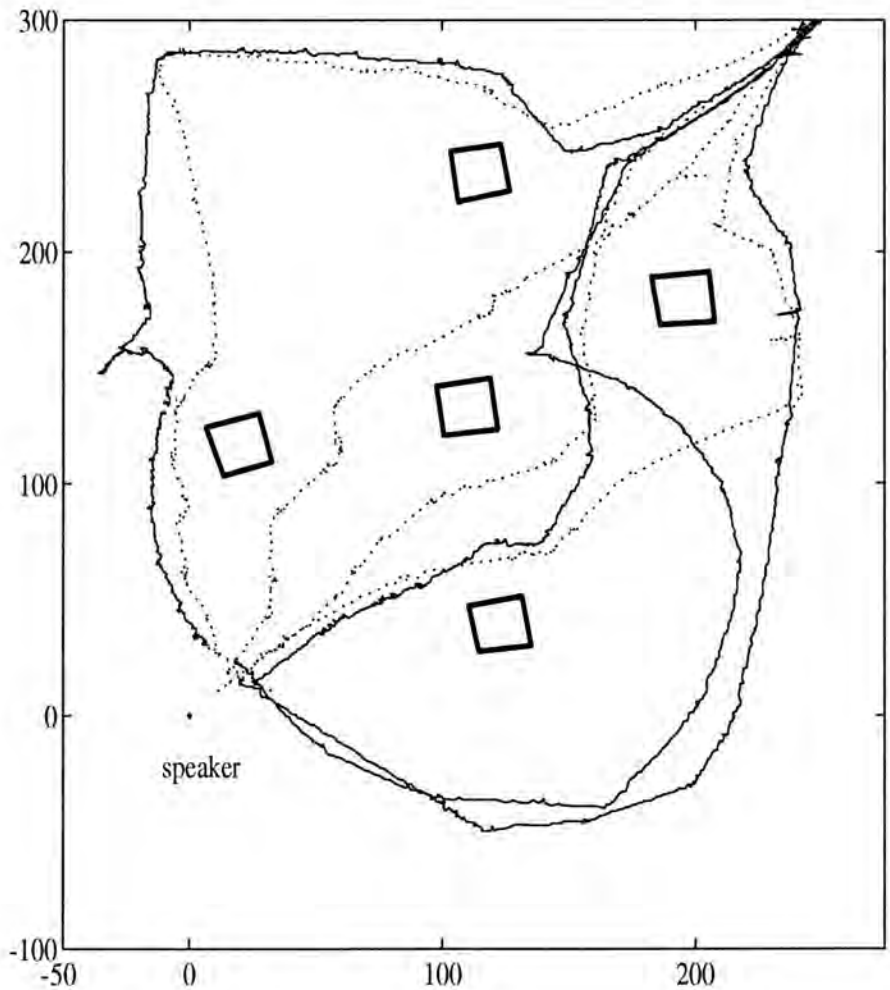



**Figure 6–15:** Eight trials with five obstacles (different line types are only to make paths clearer)

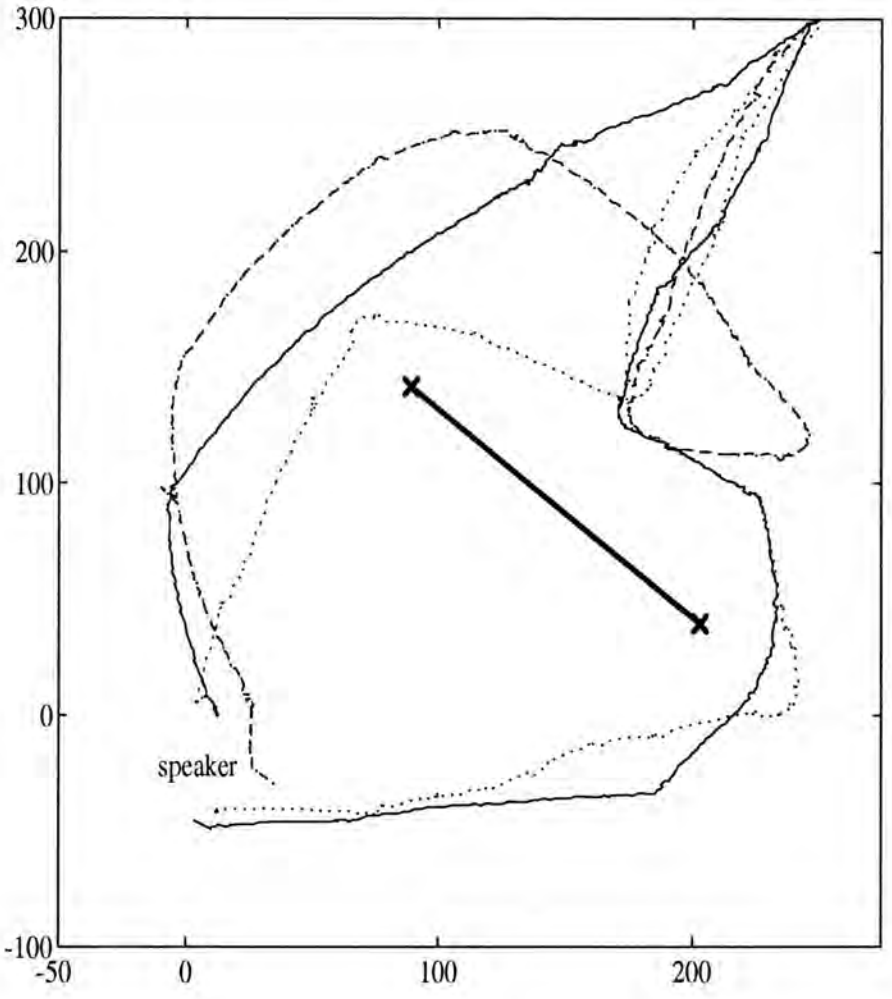



**Figure 6–16:** Five trials with a wall

eight trials with five obstacles, and five trials with a wall. The route varied but in all cases was reasonable, in that there was little backtracking or moving away from the sound. These results do indicate a reasonable degree of robustness in the mechanism. It is also worth noting that in this situation, the robot needs to be able to make substantial path corrections once it has been taken off course by an obstacle, which means that the gently curved paths associated with slower syllable rates in the open arena are likely to be less effective.

### 6.4.2 Without the arena

Some trials were also run without the arena, so that the robot could move about the whole lab. These could not be tracked so only informal observations can be made. In general, the behaviour was not as good as hoped, although the robot did find the speaker about half the time. One problem was that the robot could not always get free of the more complex obstacles (for example, chair legs) that are found in the room. Another was that the robot could be out of the effective range of the sound; particularly if it was behind the speaker. When out of range, the robot would start moving quickly which tended to reduce the chance that it would clearly hear the signal even when in range. In the latter case, having obstacles actually improved performance by forcing the robot to slow down. It would be interesting to test the behaviour in a more open acoustic environment, with less reflected sound, as this would be more like the normal cricket environment.

## 6.5 Choice

Another aspect of the natural situation that makes it more complex than the basic arena paradigm is that it is usual for groups of male crickets to sing together. They may maintain a certain distance separation from one another (about one metre) but are sufficiently close for substantial auditory overlap between songs, yet the female shows little confusion and approaches one or the other. Experimental tests have produced similar results, and these results are often taken to suggest a 'choice' mechanism in the cricket. However, as the songs are at the same frequency there is no way for them to be separately processed by the cricket's auditory system, so clearly it must be the case that the combined stimulus is still sufficient to allow tracking to be directed towards one of multiple sources.

It becomes very difficult to estimate properties of the sound field with two speakers broadcasting simultaneously: consequently it is hard to predict how the mechanism implemented in the robot will respond, or what kind of modifications

it might need to behave successfully in this situation. It seemed easiest simply to try it and see what happened.

### 6.5.1 Two songs

A second speaker was connected to the 2kHz oscillator, with timing control through a second channel on the sound-control processor, thus the two sounds could be programmed to occur together or independently. The speakers were positioned in two corners at one end of the arena, about 30cm from the corner and facing to their diagonally opposite corner. The sound level was adjusted to be about equal—difference in speaker characteristics and positional acoustics meant this was only approximate. The robot started its run from two locations, the centre of the opposite end facing towards the middle of the wall between the speakers, and from the middle of the wall between the speakers facing the centre of the opposite end. As the distance and relative angle of the speakers was very different from these two starting points the paths were analysed separately. Five trials from each starting point were run with each speaker active on its own, and ten ten trials with the two speakers producing the signal in unison. As before, the order of the runs was alternated between the different conditions (including the condition described in the next section). The signal was the same 'ideal' 1.6Hz modulation used before.

The tracks from the opposite end start point (Figure 6–17) show that with simultaneous songs, the robot goes to one speaker or the other with comparable efficiency to when that speaker alone was active. Speaker A was approached 7 times out of 10. One slight difference in the tracks is that, in the unison case, the tracks tend be more accurately directed in the lower half. This possibly reflects the behaviour demonstrated in crickets that the louder sound is usually 'chosen': once the robot has come significantly closer to one speaker, that speaker will be relatively louder.

When the robot was started between the speakers it would have to take a sharp curve to get to one speaker or the other. The resulting tracks (Figure 6–18)

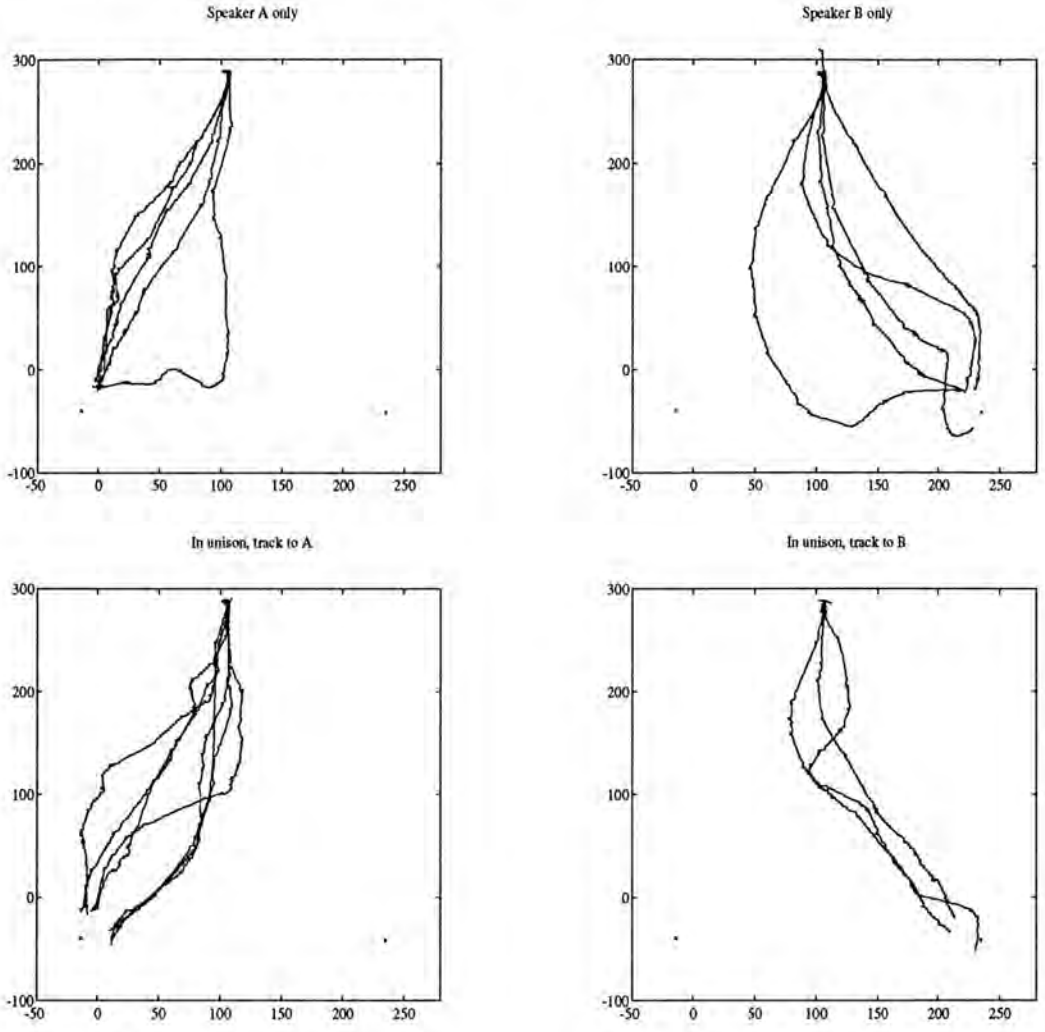



**Figure 6–17:** Tracks with speakers active individually (top) and in unison (bottom). Starting from the opposite side of the arena.

show that, with only one speaker active this movement occurs reliably; with both speakers it occurs, to one or other speaker, 6 times out of 10. In the other four cases the robot moves out of the speakers' vicinity before turning back towards one or other speaker. Thus the ability to find one sound source out of two seems a little less reliable in the situation requiring stronger corrective moves.

### 6.5.2 Variations on two songs

Studies of cricket behaviour with two songs have tended to interleave chirps from each speaker rather than present them simultaneously. However, as discussed

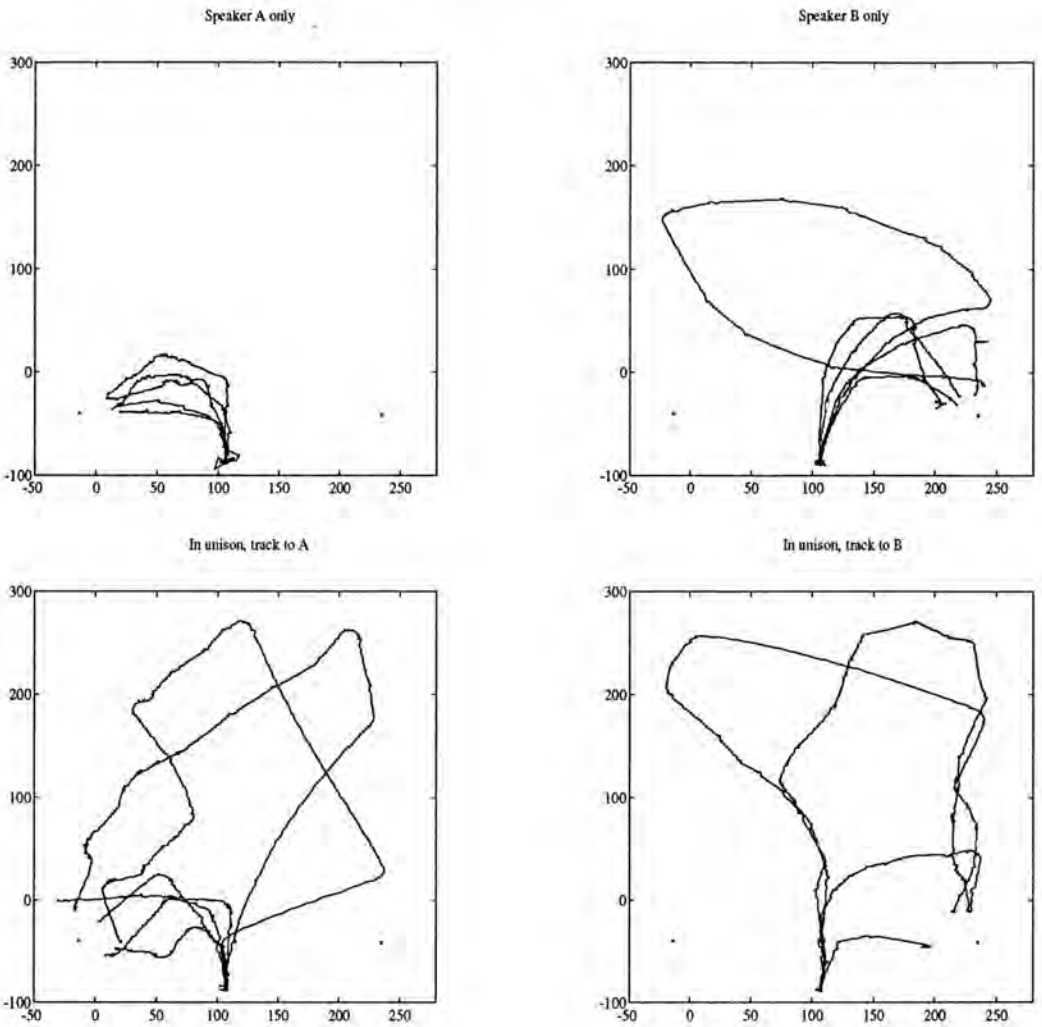



**Figure 6–18:** Tracks with speakers active individually (top) and in unison (bottom). Starting from the same side of the arena.

above, differences in the robot situation make success with this paradigm unlikely because the small arena and large movements of the robot made it difficult to show chirp response already. Weber and Thorson (1988) also report interleaving the syllables, but this creates a sound field with only small gaps between the sounds. Trials with the robot showed that it couldn't track under these conditions. These results with the cricket tend to support the suggestion earlier that the movement may involve integration over more syllables in the cricket than it does in the robot.

One experiment with the cricket could be more easily repeated here. Weber and Thorson (1988) found that with the syllables of an 'ideal' song ocurring alternately

in each speaker, so that the combined sound is a song rather than there being a complete song from each, the cricket would tend to track as though approaching a speaker located between them. Doing the same for the robot produced a somewhat similar result (Figure 6–19).

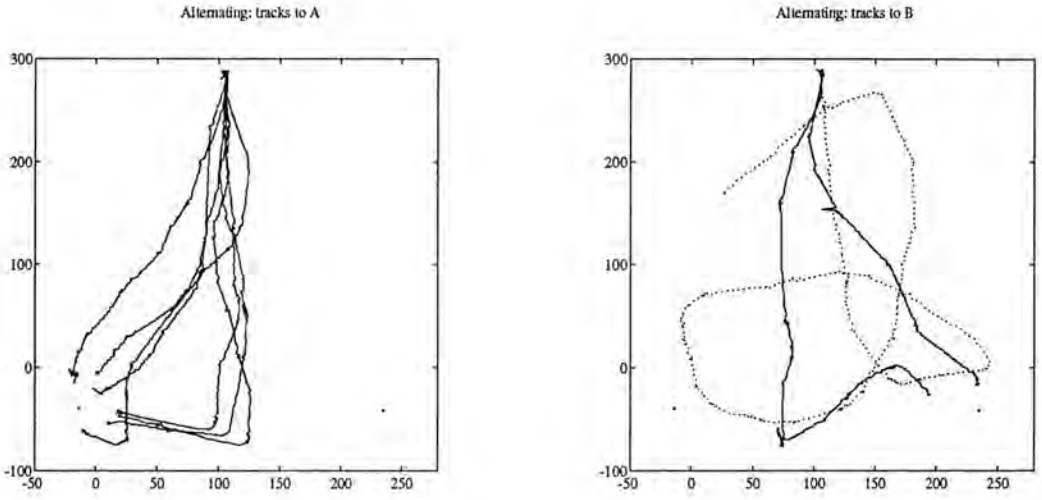



**Figure 6–19:** Tracks with alternating song

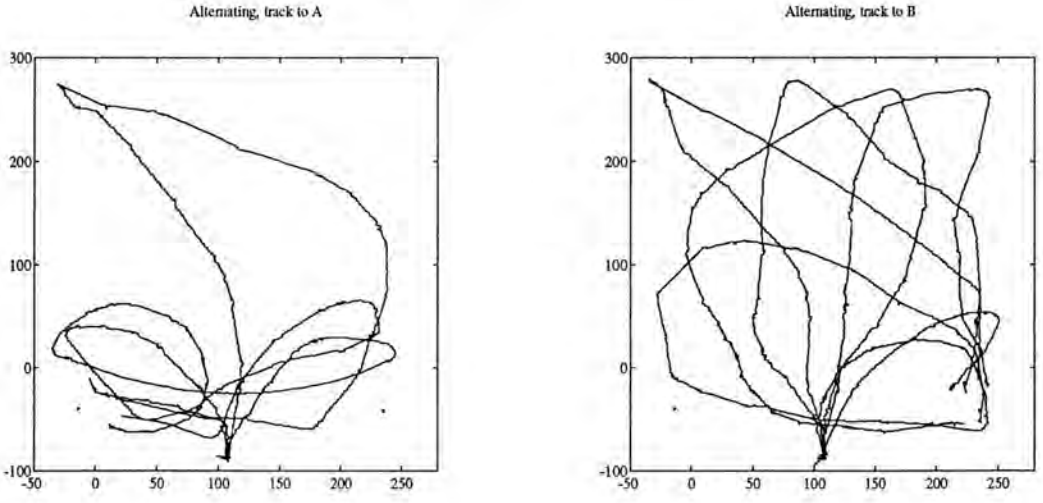



**Figure 6–20:** Tracks with alternating song, starting near speakers

Four of the tracks from the opposite wall go between the speakers, till they hit the wall and turn to one or the other, and a fifth track fails to arrive at a speaker at all. Using an origin between the two speakers to derive vector tracks for this and the previous condition (sound in unison) it can be seen that the vector means

are more scattered for alternating sound, and more closely grouped around one or other speaker position for simultaneous sound. Using distance of the means from the nearest speaker location as an ordering (as before) shows that the tracks with simultaneous sound are more direct than with alternating ($U = 20.5, p = 0.0116$, Figure 6–21).

Starting from between the speakers only two of the trials showed a direct approach, and there tended to be much wandering before arriving at one or other speaker (Figure 6–20).

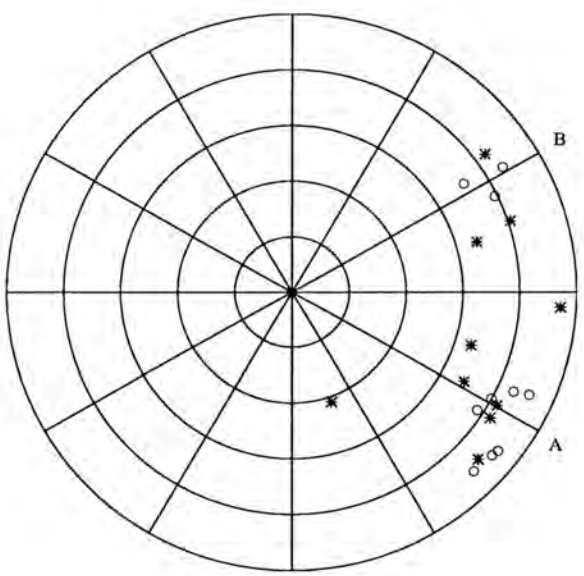



**Figure 6–21:** Mean vectors for ten trials in each condition; 'o' is in unison, '$*$' is alternating song

### 6.5.3 Two sounds and obstacles

To further test how the mechanism coped with complex situations, eight trials were run with simultaneous songs from speakers in two corners and five obstacles in the arena. As Figure 6–22 shows, in every trial the robot reached one of the

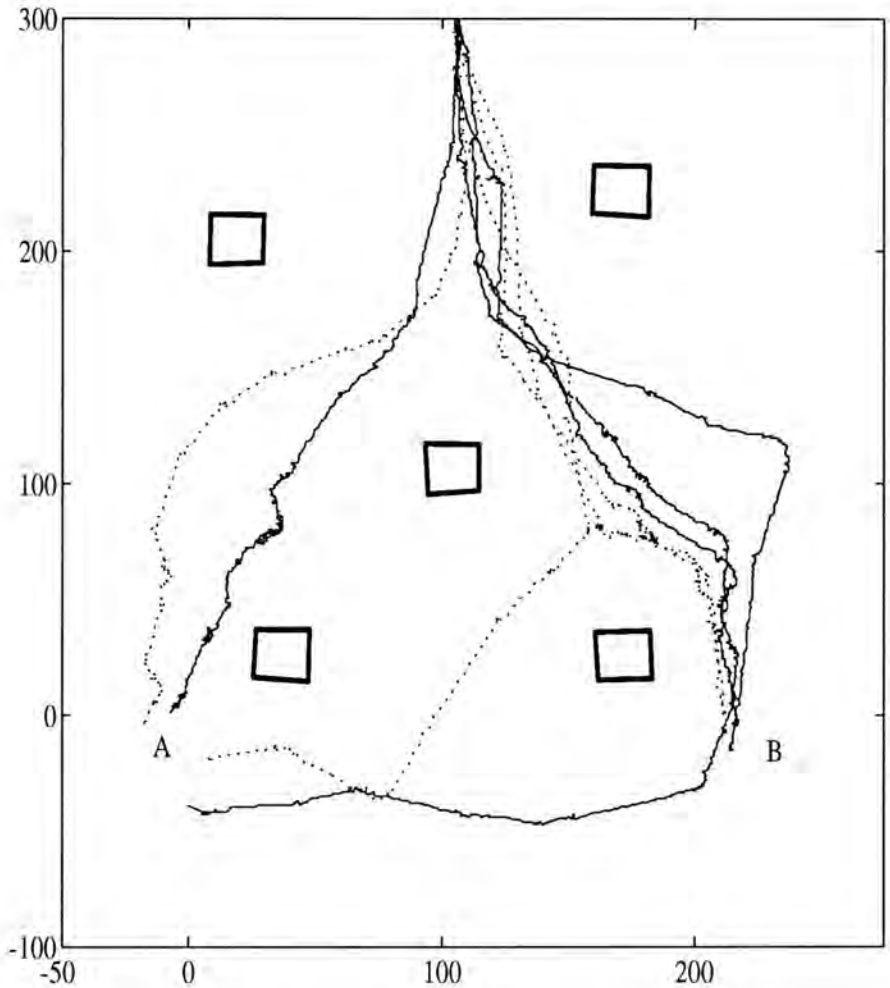



**Figure 6–22:** Eight trials with five obstacles and two sound sources

two speakers, and in general the route taken was quite reasonable.

### 6.5.4 Discussion

The success of the robot with two songs was quite surprising as it was assumed that the implemented mechanism was probably too simplistic to cope with the more complex situation. It suggests that there is no necessity to invoke cognitive attributes such as recognition, choice and preference to explain the cricket's behaviour, as it is known that no such processes are taking place in the robot when it approaches one of two sounds.

A number of more complex 'choice' experiments have been attempted with the cricket, including variations on the 'split-song' (Weber and Thorson, 1988); and testing preference for various combinations of intensity, syllable rate and chirp rate parameters (Stout *et al* 1985). The experimental equipment used in the tests here did not allow controlled variation of intensity, or adequate response to chirps, and so it was not possible to examine these more detailed results.

## 6.6 One ear taxis

A number of aspects and results of cricket behaviour have not been reproduced here. This is partly due to limitations in the experimental set-up, and partly limitations in the level of detail of the robot model: both might have been improved if more time were available. Some of these aspects have already been mentioned above: they include further 'choice' studies, greater variation of the duty cycle, a wider range of environments and clearer chirp results.

The latter results were limited by the size of the arena. Another example of this limitation was an attempt to investigate the fact that a certain proportion of crickets can still locate a sound source when one ear has been removed. Note that as the cricket auditory response at one ear is modulated by phase cancellation from the spiracle, removing an ear does not eliminate the directionality, but rather removes the input from the auditory nerve on one side. The equivalent alteration in the robot is to set one of the *response* variables at zero, rather than removing a microphone.

In one-eared taxis, crickets sometimes circle continuously, sometimes circle occasionally but with net movement to the target, and sometimes track quite well. The remaining auditory input will have a directional dependence, so it is a question of whether the comparison mechanism can operate on this well enough to produce an effective response. However, with the robot in the arena, circling almost inevitably brought the robot to a wall, which it would then turn back to avoid, and end up following. Thus it was not possible to tell whether there was

an effective angle at which the remaining *response* variable would fall low enough that it wouldn't cause a turn, allowing the robot to move forward at that angle rather than circling continuously (if the angle at which this occured were less than 90° from the sound direction, such behaviour might suffice to get to the sound eventually, albeit indirectly).

## 6.7 Summary

The following is a summary of the results presented in this chapter, the implications of which will be further discussed in the next chapter.

Like the cricket, the robot is able to find its way to a sound source using only auditory cues to guide it. The path it takes does not deviate much beyond 60° to the correct heading, and is less than twice as long as the straight line distance from the starting point to the sound. Both these factors are taken as indicators of 'direct' taxis in the cricket. There is some evidence for improved accuracy with higher intensities. Without sound the robot's movement is determined by obstacle avoidance.

The ideal sound for taxis is one with a temporal pattern of syllables at a rate of around 1.6Hz, at a frequency of 2kHz. Phonotaxis does not occur to a continuous sound of the same frequency, and is reduced outside the ideal syllable rate, in a manner that resembles the cricket's 'recognition' of a conspecific song. Syllable length is not critical, although unlike the cricket, not all duty cycles are equally effective. Tracks with chirps can sometimes be faster than those without.

When two songs occur simultaneously in different locations, the robot approaches one or the other, in a manner comparable to the 'choice' behaviour of the cricket. When the song is split between two sources, the robot shows a tendency to go between them rather than directly to one or the other, again resembling the behaviour of the cricket under similar conditions.

The cricket's natural environment provides many obstacles to a straight approach to the sound. In tests with different kinds of obstacles in the arena the

robot performed quite effectively. In the larger and more complex environment of the lab, it reached the sound source about half the time. Comparable 'success rate' measurements for the cricket are not available.

# Chapter 7

# Evaluation

## 7.1 Introduction

In the last two chapters (5 and 6) I have described the construction and testing of a robotic model of a hypothesised mechanism for cricket phonotaxis. In this chapter that model will be evaluated in three ways: as a hypothesis about phonotaxis; as an AI model of a biological system; and as an example of a functional perceptual system.

Thus I will relate the work to the three main arguments have been presented in this thesis. In Chapter 2 it was argued that essential characteristics for understanding and achieving effective perceptual systems were task-dependency, constraint exploitation, and agent-environment interfaces. In chapter 3 it was argued that using AI methods to model and test hypotheses about biological mechanisms is viable, but requires these aspects of perceptual systems to be properly represented, and thus can most usefully be done by building robotic models. In chapter 4 I argued that this approach to perception, for the particular case of cricket phonotaxis, suggests a particular mechanism for this sensory motor task that combines recognition and approach. The work has resulted in confirming evidence for each of these ideas.

## 7.2 Phonotaxis

### 7.2.1 Hypothesis

The problem of how crickets perform phonotaxis can be examined from a number of different viewpoints, and not all the questions about the mechanism that have been raised have been addressed here. Rather, the intention was to try to find the 'simplest' mechanism that could support the observed behaviour of the cricket, and to answer the question (posed explicitly in Schildberger and Horner (1988) see 4.3) of 'where and how' recognition and location may be combined in phonotaxis. The essence of the proposed solution was to attribute the lack of approach to signals that were not like the calling song to the failure of the 'location' mechanism, rather than the failure of an independent 'recognizer'. This required a 'taxis' mechanism that was inherently dependent upon frequency and temporal pattern.

In crickets, frequency dependency is implicit in the auditory receptors, partially through tuning of the receptor cells. More critically, the difference in response to sounds from different directions, necessary for taxis, is provided by physical phase cancellation, which depends upon the wavelength of the sound. Thus sound of the wrong frequency fails to affect the auditory system in a manner that enables successful location of the sound direction.

Temporal pattern dependency can be implicit if using latency differences as the basis for comparison in determining the direction to move. The critical aspect of the temporal pattern is that the syllable rate falls within a certain range. The response properties of the auditory neurons, particularly AN1, both create a latency difference based on the strength of the auditory signal, and filter out the temporal pattern (and thus the latency difference) at fast rates of change in the signal. It is worth noting that the other potential basis for comparison, the firing rate (as suggested by Huber (1988) or Doherty (1991)) is not filtered out. Comparison of latency, responding to prior onset, will preserve the temporal pattern of onsets for the stronger signal. Subsequent to the comparison, a motor response based on

summation will fail at slow rates of onset in the signal. Thus a signal slow enough to be coded by the auditory neurons, but fast enough to drive the motor response, *i.e.* that falls within a certain range of onset rates, is required for taxis to occur.

In short, the suggested mechanism for taxis involves: phase cancellation to create a directionally dependent response; a neural summation that creates a latency difference and acts as a low-pass filter; a comparison of onset time; and a neural summation for the motor response that acts as a highpass filter. While each of these components has been considered a possible part of phonotaxis by cricket researchers, this combination to explain the behaviour as a whole is, so far as I am aware, a novel hypothesis.

### 7.2.2 Is the hypothesis supported by the robot model?

Implementing a hypothesis in a model, as was done here, cannot constitute a complete evaluation of it. Evaluation involves demonstrating that the hypothesis explains the results; is consistent with known facts; and produces predictions that are confirmed. Model building is mostly concerned with showing that the results can be explained by the hypothesis, in cases where it is difficult to tell by analysis. It relies on consistency with known facts for establishing whether this explanation may be valid for the real system, but should also produce predictions for which confirmation can be sought. In other words, the purpose of building the robot is to test whether the hypothesised mechanism *can* produce the phonotactic behaviour. Applying the results to the cricket depends upon the model's consistency with known facts about the cricket mechanism. The model may additionally provide some suggestions about what further experiments on crickets would support or falsify the hypothesis.

#### Behaviour reproduced?

The results given in Chapter 6 have supplied a reasonable amount of evidence that the hypothesis is able to explain most of the fundamental aspects of behaviour that

characterise cricket phonotaxis. It has been shown that the mechanism is sufficient for approaching a sound source; that the taxis is dependent on a particular temporal pattern defined by syllable rate; that taxis can occur despite obstacles and with conflicting sound sources; and the behaviour produced is consistent with cricket behaviour in a number of respects.

There are additional elements of cricket behaviour that were not reproduced, such as taxis unaffected by duty cycle, one-ear taxis, and preferences with two different songs. Testing song preferences would require improvements in the experimental paradigm, such as better control over the sound sources; results for the duty-cycle would require refinements of the central processing; explaining one-ear taxis might require modifications to both. These behaviours are also harder to demonstrate for crickets, and vary more between subjects and paradigms, so the failure to reproduce them does not immediately contradict the possibility that the basic mechanism used in the robot is comparable to that in the cricket. Rather, it suggests further work is required to fully determine the value of the hypothesis. However, such further work probably requires knowing more about the cricket than is currently the case, *i.e.* at this stage, further evaluation is probably best done through cricket experimentation (see 'Predictions' below).

**Consistent with cricket?**

While the results do show the mechanism producing many of the appropriate behaviours, extending these conclusions to the cricket requires demonstrating that the robot is an adequate representation of the cricket, for the purpose of evaluating this hypothesis. Various aspects of the adequacy of representation have already been discussed: the description of the methods in Chapter 5 includes details of how the robot resembles or differs from the cricket. Two general observations in favour are:

- The precision of sensing and motor control is generally worse in the robot than would be expected for the cricket, so the success of the mechanism is unlikely to be due to abilities that the cricket could not really have. The

robot system is a *subset* of the cricket system rather than an *abstraction* of it.

- A number of the aspects of behaviour demonstrated *emerged* in the vehicle's behaviour rather than being specifically programmed. The robot wasn't tuned to get matching behaviour by readjusting parameters; nearly all the values were predetermined by the robot's sensory and motor capacities.

On the other hand there are some significant differences that may have influenced the results. These include:

- Using a quarter wave-length difference between the microphones possibly exaggerated the phase shift, given that in the cricket the spiracle seems to make the major contribution to phase cancellation (see 4.3.3). However, with its ears on its legs, the cricket may have a larger angle for auditory input than its body position (walking direction) indicates, consequently increasing the actual phase shift at any particular walking direction.

- Phase cancellation is mechanical in the cricket, but based on an 'ideal' equation in the robot, which means that the sources of error and noise in the two systems may differ.

- The lowpass filtering response properties of the AN1 neuron have not been sufficiently well explored in the cricket to be sure that they eliminate fast signals: there may only be a partial contribution. The implementation in the robot used a sharp cut-off for the length of gap that could be detected, which is unlikely to be the case in the neural implementation. Similarly there is a sharp change between the repetition rate sufficient for small and large turns. As it happens, neither of these was strongly evident in the results because other noise effects tended to smooth the crossover point, but nevertheless this may have contributed to the significant difference found in the response to different rates.

- Having the mechanisms of high and lowpass filtering inherently bound up in the sensory and motor properties may limit the ability of the model to explain phonotactic specificity in all cricket species. The 'recognisability' of the song serves to have females attracted only to conspecific males; more complex notions of a 'recognizer' have been based on the idea that the female has some sort of 'template' of the correct song that has either evolved with or from the mechanism that serves to generate the song in the male. Although most songs can be characterised as repetitive temporal patterning, not all preferences can be subsumed under 'bandpass filtering' (for example, *Gryllodinus kerkennenis* prefer continuous to temporally-patterned sound (Ewing, 1989)).

A significant point is that these limitations generally arose when specific decisions about how to implement the hypothesis had to be made. It is difficult to compare the plausibility of this hypothesis for crickets with alternative hypotheses because comparable specification of exactly how they could be implemented has not been made. For example, a separate mechanism for recognition might be better able to explain species variety in preferred songs; but template comparison would be hard to implement in a noisy neural system. Using relative firing rate to compare signals from each side may seem a straightforward scheme, until it is considered how the time interval chosen for comparison will be critical with a non-continuous signal.

The results given here do not show that comparison could not be based on firing rate rather than latency; nor that a separate mechanism for recognition would not be effective. However, from the discussion in 4.4, it is arguable that no *simpler* model can be described that could potentially encompass the same range of behaviours, and still be effectively implementable in a mobile robot.

### Predictions?

Building and observing the mechanism in the robot did suggest a number of experiments for crickets that might help determine whether the hypothesis is a viable

one. Carrying out such experiments may be difficult, some more so than others, but the hypothesis does indicate some interesting directions for investigation.

One noticeable aspect of the robot behaviour was that the modes of failure of taxis reflected the functioning of the mechanism: taxis was inefficient rather than non-existent at slow rates, for example. Cricket studies in general have not specified in what manner the cricket goes wrong when it does go wrong; it may be possible to determine whether the failure modes are consistent with failure of the taxis mechanism rather than a lack of recognition. Robot behaviour also suggested a possible adaptive advantage to the chirp structure of song that explains why it may improve taxis (act as a 'motivational parameter' according to Popov and Shuvalov (1977)) while not being necessary (not an 'essential recognition parameter') for taxis to occur. Examination of the response of crickets to trills and chirps under different paradigms may bear out this suggestion. In general, a more careful consideration of *what exactly* has improved when the cricket does 'better' taxis seems warranted—reporting 'percentage of time tracking' (as do Thorson *et al*, 1982) is less useful than giving values for each of the criteria applied to determine tracking (see 6.2).

The ability to interpret results for complex sound fields is currently limited by the inability to estimate the tympanal response to such situations (Weber and Thorson, 1988). Measuring the tympanal, or auditory nerve, responses in these situations would certainly make reasoning about the taxis mechanism on the basis of these experiments more viable. It may also be possible to supply the cricket with tympanal or auditory nerve stimulation that separates the latency and intensity, and thus directly examine how each contributes to the turning response—experiments of this kind have been carried out on the grasshopper (Romer and Rheinlaender, 1983).

Further examination of the response properties of brain neurons seems warranted. There does not appear to have been any replication of the widely quoted results of Schildberger (1984). In particular, the hypothesis presented here suggests that their high-pass selectivity may be based on the compared signal, so the response properties with sound from different locations should reflect the direc-

tion. Further, this response property has been associated with requirements of the motor mechanism, so determining details of the further connections of these neurons would be interesting, although possibly rather difficult to do.

### 7.2.3 Extensions

Although the model is consistent with cricket research, it is also a simplification of it. There are several ways in which it might be extended or improved: ranging from slight elaborations in the program to complete reconstructions of the robot. Some of those possibilities will be discussed here.

**Experimental procedures:** There were several limitations on the experimental set up that limited the results. A larger area to move around and an acoustically open environment would be useful. Improving the quality of the sound source might also help in investigating more complex capabilities.

**Neural details:** The current model reflects few details of the known functional properties of the neurons. A more biologically plausible implementation of these processes, would improve the credibility as a hypothesis of neural functioning. It might also help explain some of the anomalies between results in neurology and behaviour, such as taxis with short syllable gaps. The potential role of some of the additional auditory neurons, especially ON1, could also be explored.

**Improved sound processing:** The circuit for processing the sound is less than ideal in a number of respects. One major improvement would be to increase the sampling rate to the order of 1kHz so that it is more comparable with the response rate in the cricket, and not so slow compared to the turning speed in the robot. Even more appropriate would be to build an ear-and-tube that performed physical phase cancellation. While this would take some tuning, it would combine amplitude and phase effects in the same manner as the cricket and thus be a better model of noise and error conditions,

although replicating exactly the acoustic characteristics of the cricket ear is not currently possible.

**Additional sensors and behaviours:** The sensing for obstacle avoidance is quite minimal, certainly orders of magnitude below the complexity of visual and tactile sensing of the cricket. Getting reliable behaviour from the robot in a wider range of enviroments would need at least some improvement of these sensors, for example to allow it to find its way through smaller gaps , or more efficiently avoid getting caught on awkward obstacles. Another reason to provide additional sensory systems would be to implement a better model of how phonotaxis interacts with the other natural behaviours of the cricket. This might be significant in how the behaviour changes when the signal for phonotaxis is less than ideal.

**Legged locomotion:** It seems likely that the mechanism for taxis is to some degree determined by the requirements for controlling turning in walking, which will differ from those for wheeled motion. Implementing the mechanism on a six legged robot would be useful to explore how significantly it does affect the design. This might also permit the possibility of addressing problems such as taxis over rough terrain.

**Reduced scale:** Some of the problems in drawing comparisons with the cricket were due to the difference in size. Recent technological developments for building small walking machines make it possible to speculate about building a robot model of a size more comparable to the cricket. Specialised circuitry to implement the mechanism could fit on a microchip. Such a robot could be placed in identical situations to the cricket for direct behavioural comparisons.

## 7.3 AI and Biology

There are many means of investigating hypotheses about biological mechanisms. Building models of sensory motor-systems in AI is a relatively recent approach, and its methodology is not well established. Modeling can potentially make a number of contributions to our understanding of perceptual systems. I will illustrate how model-building has aided explanation in this particular case. However, as I argued in Chapter 3, the contribution depends upon the *kind* of model used. Again I will assess how the specific approach taken to modeling in this work has affected the outcome.

### 7.3.1 Usefulness

#### Downhill synthesis

The principle that Braitenburg (1984) calls 'downhill synthesis'—that it is easier to put together a mechanism that produces apparently complex behaviour than it is to work out from the behaviour how it was produced—can be an advantage in approaching a problem, such as phonotaxis, from the perspective of trying to build a device that does it. Phonotactic behaviour can appear quite complex in the cricket, but this does not necessarily imply a complex mechanism underlies it.

The hypothesis tested here was substantially simpler than those that have been proposed by researchers analyzing cricket behaviour. Planning to build a mechanism provides a strong impetus to look for simpler means to achieve the required functions. With a patterned signal, latency comparison appeared easier to implement than 'firing rate' comparison; and as this latency response would be affected by the temporal pattern, it thus seemed an appropriate way to incorporate the 'recognition' effects. The behaviour of the robot does turn out to appear more complex than might be expected from the internal mechanism, in illustration of the principle above. This is most noticable in the case of two sound sources: the

robot appears to choose one and ignore the other, but in fact there is no possibility that the two sounds are being separately processed.

### Completeness

Another effect of building working models of hypotheses is that it is not possible to leave parts of the mechanism unexplained. A reasonably complete theory is necessary, although some elements may be more conjectural than others. Also, requiring actual circuitry and programming code prevents any vagueness in the specification of the parts.

This requirement relates to earlier discussion of the importance of approaching perceptual systems as a complete linkage from sensors to actuators. Building a working robot means that this approach *must* be adopted: considering sensory processing in isolation is not possible. Trying to do this for phonotaxis reveals that there has been almost no attempt in research on crickets to actually trace the processing from sensors through to motor control.

In fact, building a model reveals, on a number of levels, areas where current research is insufficient. Another example is the difficulty of incorporating the known properties of the AN neuron with the behavioural results. Other examples have been given above. It also reveals the fact that reported results are often not sufficiently specific for their significance, with respect to a hypothesis, to be determined.

### Validation

A working model is quite a convincing demonstration that a hypothesised mechanism does account adequately for the behaviour. Theoretical descriptions of mechanisms are often difficult to evaluate, especially if not expressed formally. Moreover, the process of getting that mechanism to work can reveal strengths and weaknesses of the hypothesis that may not be obvious from a verbal statement of it. For example, will it require many ad hoc adjustments? While the hypothesis regarding taxis presented here can be supported by examining evidence from

crickets, it is certainly strengthened by the fact that it was possible to build into a robot, which then produced comparable behaviour.

Another advantage is that it may be possible to show what a certain mechanism will do under various conditions despite being unable to formally analyze the effect of those conditions. An example here was the use of two sound sources: it would be difficult to estimate the effect on the sensors well enough to prove in advance whether the mechanism would work; but by putting the model in the situation it can be determined whether or not it *does* work. As many sensory-motor devices are required to work in complex sensory conditions, such a 'try-it-and-see' approach is perhaps more generally useful than attempting to find and solve the physical equations that describe the situation, although explanation of the success or failure will still require analysis.

### 7.3.2 Robots vs. Computer Simulation

At least some of the advantages discussed above would apply to an attempt to simulate computationally, rather than physically, the system of cricket phonotaxis. The impetus for a simpler mechanism, and the need to specify it in sufficient detail to express it in computer code, would still apply. In other areas some advantages could be lost: 'completeness' would depend upon the detail of the modeling of the environmental situation; and validation would be less direct. Parts of the problem might be missed because of inadequate representation.

There are further specific reasons why computer simulation of this behaviour would probably provide weaker support for the hypothesis than that provided by the robot. One is that taxis, approach along a sensory gradient, does not require a complex behavioural solution: the principle of turning to the side more strongly stimulated is fairly obvious, and it is not necessary to model it to accept that it will work. A simulation of taxis is only interesting to the extent that it tackles a real problem of detecting the differences in response to a sensory source, and integrating the response with the constraints on mobility. If the sensory detection and motor response are represented as reliable and regular (as in the simulation

described in 3.4) then achieving taxis becomes fairly trivial. In passing it should be noted that 'learning' taxis in such a situation (for example in Pierce and Kuipers (1991)) is probably not particularly difficult either.

But a computer modeling of the real physics of phonotaxis in a moving animal or robot is quite difficult. 'Ideal' sound propagation can be described by fairly simple equations, but in any real situation, with a directed speaker, a floor surface, reflecting walls and so on, it becomes extremely difficult to calculate with any accuracy. Likewise, the real physics of a motor response involve more interactions of forces than can be viably captured without complex equations (for wheeled robots see, for example, Alexander and Maddocks (1989)). In short, it would require a great deal of effort to build a computer model that reflected the real situation well enough to make strong claims that the mechanism actually works. And it was certainly the case here that a simpler simulation was quite misleading about how the mechanism would perform. Furthermore, extending the experiment to add more obstacles or sounds would have required even further elaboration of an already complex model.

Using a computer model rather than a robot would, in this case, substantially weaken the justification for extending the results of testing the hypothesis to the cricket. For example, it would be hard to claim that, insofar as sensors and motors differ, those in the model are substantially *worse* in accuracy than in the cricket: it would be more likely that the mechanism works only because of the idealised conditions. There would also be a lot more tuning of variables involved, which could well lead to values being specifically chosen to generate 'cricket-like' behaviour, rather than that behaviour resulting from getting the mechanism to work at all.

This does not demonstrate that simulation is necessarily 'bad'. But it certainly seems the case here that an *adequate* simulation would in fact be more difficult to implement than building a robot. It does suggest that for perceptual problems such as this, where sensing and motor response are tightly linked and the mechanism depends upon their specific properties, computer simulation is not necessarily the

best first step in model-building—a physical robot may more easily capture the critical aspects of the problem.

### 7.3.3 Viability

As noted in Chapter 3, it is often argued that simplistic simulation is being done because it is not yet 'viable' to build physical models of perceptual systems to test specific hypotheses. The work reported here shows that it is viable, and there seems no reason why other examples of perceptual systems could not also be approached in this way. However the usefulness of this model was dependent upon appreciation of the critical aspects of modeling discussed in Chapter 3: adhering to the following ideas permitted useful application of the results of model building to the biological system:

#### Representing appropriate aspects of the problem

Taxis is a mechanism for finding a sensory source when it is not possible to simply set a heading and go towards it; so it should be modeled by a system in which this is not possible. How it is implemented depends upon the nature of the sensory signal, so the same sensory mode and comparable sensors should be used. Real time interaction of the sensory and motor processes is an important determinant of the behaviour and should be maintained as far as possible.

#### Justifying design assumptions

Assumptions will have to be made in any modeling attempt, but it should be possible to justify those assumptions on the basis of the hypothesis and additional knowledge about the subject. In building this model I attempted to explicitly avoid internal processes that contradicted known results for the cricket; and design limitations were explicitly noted, so that their effect on the results could be appreciated (as in the discussion in 7.2.2)

**Using a framework that permits experimental evaluation**

An important aspect of the work presented here was that it was possible to make some direct comparisons between the model and the cricket. This was assisted by making experimental design considerations at an early stage; by choosing the level of detail of the model to make repeated testing possible (detailed models that takes days to complete a trial can be less informative than simpler ones); by devising appropriate means to measure the behaviour; and by having clear hypotheses that could be answered using statistical information.

### 7.3.4 Advantages of biological inspiration

Achieving these attributes in the model was facilitated by the choice of the system to be modeled. Having a well-defined target system is extremely useful in guiding decisions about representation, design and experimentation. The lack of such targets is one of the main obstacles to evaluation of the modelling of sensory-motor systems in AI. While it is true that many perceptual mechanisms in animals would be very difficult to model, this explanation of why 'biology inspired' AI has, in general, such a weak connection with actual biological systems is not justified. There are many examples provided by neuroethology that could well be explored by these methods. Taking such low-level systems as targets has a number of advantages. The problem of cricket phonotaxis involves a complete linkage of sensory to motor processes, yet is sufficiently simple that it is possible to derive informed hypotheses about the mechanism, at the level of neural processes. Thus analogy between it and the model is at a useful level for evaluating mechanisms.

Assuming that AI can simply use biology as a source of mechanisms, as seems to be advocated by Beer (1990), is unjustified however—there aren't many biological perceptual systems that are well enough understood to be directly implemented. The principal idea behind the methodology presented here is that the *process* of attempting to implement physical models of biological systems can potentially contribute to our *understanding* of how perceptual systems work.

## 7.4 Perceptual Systems

In addition to the value it may have as a model of phonotaxis in the cricket, the robot described here is a working implementation of a perceptual mechanism. As such, it can be compared to other work in the field of mobile robotics. Phonotaxis in a robot has not before been reported in any detail, as was discussed in 5.2, and reliable location of a particular sensory source using taxis, as demonstrated here, is a behaviour of comparable competence to many more complex mobile robotic systems.

In this section I will examine to what extent the successful development and deployment of this system depended upon the characteristics that were ascribed to perceptual devices in Chapter 2. In other words, does the work here bear out the value of taking such an approach to sensory systems? This has already been discussed in the context of developing a hypothesis about the cricket, here it will be applied to the workings of the robot.

### 7.4.1 Specific Task

The mechanism in the robot was highly dependent upon the specifity of the task: to approach a particular sound source. First, requiring taxis rather than localization made the requirements much simpler. The robot need only detect and respond to the sign of the difference rather than use the size of that difference to estimate position. Even compared to a localization task such as 'turn to face the sound then move directly to it', which could potentially operate using sign differences (*i.e.* turn till there is no difference) taxis is easier to implement: constant adjustment can overcome the effects of noise; the front-back ambiguity of determining direction with two inputs doesn't matter because moving away is an unstable point; and directional error of up to 90 degrees will still suffice to have the robot reach the target.

The specificity of the sound source was also essential to the functioning of the mechanism. The use of phase cancellation to detect interaural differences works because the wavelength of the sound is fixed. The simplicity of latency comparison (responding to first onset) depends upon the temporal variation in the signal. It is interesting to note that in this system, specificity is a feature rather than a bug. That is, the fact that the mechanism only works to particular inputs, in addition to simplifying the mechanism, is part of the characterisation of the task in the first place: it is not supposed to work for the wrong signals. This points to the counterproductivity of the common tendency to analyze a specific task as the intersection of several more general abilities (in this case, the ability to recognize a pattern, and the ability to approach a sound), when general abilities are actually *more* difficult to implement than specific ones.

The robustness of the mechanism, coping with obstacles, background noise and reflections, and even the two sound situation, despite the variability of the motor response with floor surface and battery power, illustrates another misconception related to that above—mechanisms don't need to be general purpose to cope with variable environments. Many robots that incorporate more complex sensory processing than this one, such as visual systems, use only a fraction of that information in the tasks they actually perform—they don't need to recognize a tree to avoid it.

### 7.4.2 No representation

This robot operates without any attempt to build an internal model of its environment: there is no centralised representation of the sensory situation, not even in a distributed sense (a non-iconic map at some layer of a network). This does not prove that such representation is never needed, but rather shows that it is not, as has been claimed, an essential component for successful sensory-motor behaviour (for example, Kreigman *et al* (1987) begin their paper "A mobile robot architecture must include sensing, planning and locomotion which are tied together by a model or map of the world ...").

It could be argued that the robot does contain 'representations', in variables such as *response* that correspond to sensory inputs, or *motorstate* to motor outputs. But does conceiving of these variables as 'representations of the external world' and thus the mechanism as 'manipulation of symbols' actually provide an explanatory function? The variables *are* representations, of the corresponding neural processes in the cricket: should these themselves then be characterised as 'symbolic'? In neither case is it necessary to use this symbolic interpretation to explain how the system functions: the variables served a mechanical function in connecting sensors to motors, a role epistemologically comparable to the function of the gears connecting the motors to the wheels.

Nor does the functioning of the robot fit usefully into the sense-think-act cycle of traditional behaviour decomposition. The robot is continuously moving and thus changing its sensory input. The important characteristics of the sensory situation are non-static: a temporal pattern can't be detected from a snapshot of the environment. The processes of peripheral sensing (in the ear and IR circuits), the central processes, and the movement are all occuring at the same time, and their dynamic interaction forms the basis of the behaviour.

### 7.4.3 Exploiting constraints

Constraints are partly determined by the task, so many of the aspects noted in the 'task' section, such as taking advantage of a single frequency, are examples of constraint exploitation, or 'matched filters'. A different example of exploitation of an environment-determined constraint was the use of independent control of each wheel because the sound couldn't be closer to both sides at once; and the simple environment of the arena without obstacles also has the property that the direction of sound is always be away from a wall, so approaching sound and avoiding walls will not interfere. The original version of the program had to be modified when this constraint was violated by introducing obstacles between the robot and the sound.

There are also 'unintentionally' exploited constraints, such as having a flat floor surface and consequently being able to use a simple wheeled robot, and having easily detected obstacles and hence using simple forms of obstacle detection. In these cases it becomes difficult to separate designing the robot to fit the environment, and designing the environment to fit the robot. However, the same can be said of biological systems—their environment evolves as they do, rather than being a fixed situation which they evolve to fit more 'optimally'. In the context of robot building, modifying or limiting the environment in some way so as to simplify the processing may be viewed as undesirable, if often necessary; but few useful machines of any kind have been designed that have *no* environmental limits, even if that is as basic as 'keep it out of the rain'.

### 7.4.4 Importance of interface to environment

The exploitation of constraints often involves having the right kind of sensory and/or motor interface to the environment. The above discussion, and the description of the architecture of the robot in Chapter 5, reveal how dependent on the details of this interface the mechanism was.

One aspect of this was the fact that morphological changes to the robot were often important to obtaining the required function. Examples include: turning the microphones to face backwards which improved the phase difference response (and of course the distance between the microphones was an essential determinant of the response); using a ball-bearing castor rather than a fixed wheel so that movement was more immediately responsive to the motors, and turning by stopping a wheel was possible; and positioning the IR sensors so that there was little overlap in their angle of view and thus minimal interference.

Further, the parameters in the program were largely determined by the properties of the sensors and motors. The rate and accuracy of the auditory response set a limit on processing speed, and the rate of turning and moving set a limit on response time. Even factors such as the amplitude of motor related noise detected by the auditory circuit affected the choice of parameters for summation rates.

But the effect is more fundamental than this. Perhaps the abstract principle of the mechanism 'Use the difference between the ears to control turning' can be stated without reference to how the sensory difference arises, or how the turning is controlled. But to be any more specific about the mechanism, these details need to be known: in what form, and at what rate is the difference available, and how, and with what accuracy does it reflect the direction of the robot; what form and rate of signalling do the motors require and how is the position affected? It is hard to identify any part of the control code in the robot that is not influenced by considerations of this kind.

### 7.4.5 Non-adaptive

This robot doesn't use a neural net, genetic algorithm, or other learning mechanism to modify its behaviour from experience. On the other hand, the development process did involve a certain amount of tuning to get the mechanism to work effectively. It might be possible to usefully incorporate some adaptive features into the model to tune it more precisely to the environment, but this would still be with the basic framework of a mechanism largely fixed in advance. The lack of self-adaptive components in this system has not led to 'brittleness', however. The mechanism was quite robust in a variable and noisy environment, and even coped with such major alterations as having a second sound source. Coping with environmental variety does not necessarily require modifiable mechanisms.

Not having tried to produce the behaviour using neural networks, for example, it cannot be claimed that doing so would not produce an effective mechanism. However, judging by current capabilities, it doesn't seem likely that they could do so within a reasonable time span. In particular, while it might be feasible to train a net to recognise the syllable rate, or to generate appropriate movements based on the input difference, it is less likely that such training could come up with the principle of using latency comparison to incorporate recognition: though advocates of the approach might argue that it could come up with something better!

# Chapter 8

# Conclusion

## 8.1 Perception and AI

This thesis set out to explore how AI methods might be applied to understanding perceptual systems. Concern with perception—connecting systems to the real world through sensors and actuators—is becoming increasingly important in AI, as it is realised that a major limitation on current systems is their dependence on human intervention and interpretation (Harnad, 1990). Perception is seen as a means to make systems more robust, and able to take into account changing conditions. It is also considered necessary to give the computational processes *meaning* (Lakoff, 1988). What is lacking is often characterised as the ability to interact autonomously with a complex dynamic world.

Such concerns have led to an increased appreciation of such 'adaptive behaviour' in lower animals, and a research movement based on the idea that getting 'life-like' behaviour from AI systems is an appropriate goal (Langton, 1989). It has been suggested that it should be possible to establish a useful dialogue between biology and AI, comparable to the interaction between psychology and AI that has spawned cognitive science (Bedau, 1992). However there have been few attempts to systematically discuss how that interaction can best be established, or what methodologies might be useful to achieve it.

## 8.2 Theoretical Contributions

Addressing these issues required an examination of the nature of the target systems in biology, and current examples of working perceptual systems for robots in AI. From this, it was argued that discussing perception in terms of mechanisms by which sensors contribute to behavioural competence is more appropriate than treating perception as a process of creating internal representations. This leads to an appreciation of the extent to which the specific task, and the physical interface between the system and its environment, shape the mechanisms that underlie perception.

Also required was an analysis of the role of AI in investigating such mechanisms. Understanding the function of model building provides reasons for advocating a methodology that differs from most current work on sensorimotor control in 'autonomous agents'—rather than supporting computer simulation of adaptive mechanisms for generic creatures, I contend that a deeper understanding of perception requires physical modeling of specific biological systems, to generate and test hypotheses about how they function.

## 8.3 Practical Contributions

The problems approached in the work in this thesis lay on a number of levels. At the lowest level, there was the problem of building a functional robot, including the physical architecture, electronic circuits for sensors, electronic-processor interface, and control programming. The resulting robot displayed an interesting level of competence at a sensory motor task that has not, to my knowledge, been previously attempted—that of locating a particular sound source under varying environmental conditions.

Another important aspect at this level was the development of experimental procedures and methods for measuring the robot's behaviour. Substantial analy-

sis, involving specially designed data processing programs, was carried out. This was more useful for testing hypotheses than the comparable levels of processing required to produce picturesque simulation graphics—which can be useful but can also be misleading.

The development of these low-level systems interacted in several ways with the next level—the testing of a specific hypothesis about phonotaxis in the cricket. Designing the robot and establishing measurement procedures were facilitated by having a clear target system to address. Building these systems underpinned the ability to draw some clear conclusions about the viability of the hypothesis. Deriving the hypothesis itself was aided by the considerations raised in attempting to build a robotic model.

At a higher level, the hypothesis also emerged from a particular characterisation of the perceptual problem of phonotaxis. Approaching perception by considering the behaviour it supports, and the environment-agent interaction it involves, has been shown to have direct application, both for understanding particular systems and also as a background for methodological decisions. The work demonstrated that it is possible to have a close, useful connection between biology and AI, at the level of testing hypotheses about the neural mechanisms of perception.

## 8.4 Future directions

In 7.2.3 I have discussed some specific extensions to the work done on the phonotaxis mechanism. Many of those are directed at improving the accuracy with which the robot models the cricket. The current model was aimed at testing a fairly basic hypothesis; establishing a more detailed explanation of how phonotaxis is controlled by the cricket's physical and neural structure would require that structure to be more adequately represented.

Specifically, increasing the similarity of interaction with the environment would be an important first step: improving the auditory circuit to produce faster readings, and working towards a more appropriate locomotion system. Improving the

experimental set up also would permit closer comparisons to cricket behaviour, and make it potentially possible to derive specific predictions for the cricket on the basis of trials with the robot. However, it would probably be necessary gather more data on the cricket itself before making decisions about improvements to the robot's control program.

Another direction for further research would be to attempt to implement alternative mechanisms hypothesised to explain phonotaxis. As I have already discussed, there are none that have been explicitly presented, so it would be necessary to derive such hypotheses first. Some possibilities include using relative firing rates rather than latencies (although there would then need to be an alternative means for recognizing the song); using temporal correlation, or template matching, rather than summation to filter out incorrect signals; or having the filtering occur before the two sides are compared, so that the 'strength of recognition' is the basis for which way to turn. Comparing the relative difficulty of implementing such mechanisms, and their relative performance under the experimental paradigms used here, would provide a background for stronger conclusions about the value of the particular hypothesis I have presented.

More generally, the potential for adopting similar methods to examine other examples of sensory-motor behaviour in biological systems seems promising. The work of Franceschini *at al* (1991) with a robot model of fly vision is the only directly comparable research of which I am aware, but there is sufficient biological information and robot-building experience that many similar investigations could be undertaken, to the benefit of understanding perceptual systems from both perspectives. To pick just two examples, the cockroach escape response described by Camhi (1988), and the mantis shrimp's visually guided striking action (Di Stefano *et al*, 1990), have both been described at a level of detail that would make robotic modeling possible.

## 8.5 Artificial Insects

It might be questioned whether substantial work to create a detailed model of one specific sensory-motor mechanism will contribute to a more general understanding of the functioning of perceptual systems. This is difficult to answer: it has already been argued that close examination of biological systems gives the impression of a multiplicity of specialised mechanisms, each appropriate to the environmental niche of a particular species. As McFarland (1991) has argued "Just as there are no general-purpose animals, so there should be no general-purpose robots". There are likely to be some common mechanisms that evolve under similar pressures, but there is unlikely to be one simple set of principles that are the key to the production of successful behaviour. Rather, the study of perception requires an appreciation of the complex and varied structures that underlie it. This can only be gained by in depth examination of real systems that use sensing to control action.

# Bibliography

[Alexander & Maddocks, 1989] Alexander, J.C. and Maddocks, J.H. (1989). On the kinematics of wheeled mobile robots. *The International Journal of Robotics Research*, 8:15–27.

[Anderson & Donath, 1990] Anderson, T. and Donath, M. (1990). Animal behaviour as a paradigm for developing robot autonomy. *Robotics and Autonomous Systems*, 6.

[Arbib, 1987] Arbib, M.A. (1987). Levels of modeling of mechanisms of visually guided behaviour. *Behavioural and Brain Sciences*, 10:407–465.

[Arbib, 1989] Arbib, M.A. (1989). Interacting subsystems for depth perception and detour behaviour. In Arbib, M.A. and Amari, Shun Ichi, (eds.), *Dynamic Interactions in Neural Networks: Models and Data.* Springer-Verlag, New York.

[Arbib & House, 1987] Arbib, M.A. and House, D.H. (1987). Depth and detours: an essay on visually guided behaviour. In Arbib, M.A. and Hanson, R.A., (eds.), *Vision, brain and cooperative computation.* MIT Press/Bradford Books, Cambridge, Mass.

[Arkin, 1989] Arkin, R.C. (1989). Motor schema-based mobile robot navigation. *The International Journal of Robotics Research*, 8:92–112.

[Arkin, 1991] Arkin, R.C. (1991). Integrating behavioural, perceptual, and world knowledge in reactive navigation. In Maes, Pattie, (ed.), *Designing Autonomous Agents* MIT Press.

[Atkins *et al*, 1984] Atkins, G., Ligmann, S., Burghardt, F. and Stout, J.F. (1984). Changes in phonotaxis by the female cricket *Acheta domesticus L.* after killing indentified interneurons. *Journal of Comparative Physiology A*, 154:795–804.

[Bailey & Thompson, 1977] Bailey, W.J. and Thompson, P. (1977). Acoustic orientation in the cricket *Teleogryllus Oceanicus (Le Guillou). Journal of Experimental Biology*, 67:61–75.

[Bailey *et al*, 1990] Bailey, W.J., Cunningham, R.J. and Lebel, L. (1990). Song power, spectral distribution and female phonotaxis in the bush cricket *Requena verticalis (Tettigonidae: Orthorptera). Animal Behaviour*, 40:33–42.

[Ballard, 1991] Ballard, D.H. (1991). Animate vision. *Artificial Intelligence*, 48:57–86.

[Bedau, 1992] Bedau, M.A.. (1992). Philosophical Aspects of Artificial Life. In *Proceedings of the First European Conference on Artificial Life*, Cambridge, Mass. MIT Press.

[Beer, 1990] Beer, R.D. (1990). *Intelligence as Adaptive Behaviour.* Academic Press, San Diego.

[Beer & Gallagher, 1991] Beer, R.D. and Gallagher, J.C. (1991). Evolving dynamical neural networks for adaptive behaviour. *Journal of Adaptive Behaviour*, 1.

[Beer *et al*, 1990] Beer, R.D., Chiel, H.J. and Sterling, L.S. (1990). A biological perspective on autonomous agent design. *Robotics and Autonomous Systems*, 6:169–186.

[Beer *et al*, 1992] Beer, R.D., Chiel, H.J., Quinn, R.D. and Larsson, P. (1992). A distributed neural network architecture for hexapod robot locomotion. *Neural Computation*, 4:356–365.

[Belew, 1991] Belew, R.K. (1991). Artificial Life: A constructive lower bound for Artificial Intelligence. *IEEE Expert*, 6:8–15.

[Bourgine & Varela, 1992] Bourgine, P. and Varela, F.J. (1992). Towards a practice of autonomous systems. In *Proceedings of the First European Conference on Artificial Life*, Cambridge, Mass. MIT Press.

[Boyd & Lewis, 1983] Boyd, P. and Lewis, B. (1983). Peripheral auditory directionality in the cricket. *Journal of Comparative Physiology A*, 153:523–532.

[Braddick, 1980] Braddick, O.J. (1980). Direct perception: an opponent and a precursor of computational theories. *Behavioural and Brain Sciences*, 3:381–382.

[Braitenberg, 1984] Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology.* M.I.T. Press, Cambridge MA.

[Brentano, 1874] Brentano, F. (1874). Psychology from an empirical standpoint. In Sahakian, W.S., (ed.), *History of Psychology.* F.E. Peacock Publishers, Itasca, Ill. (1981)

[Brooks, 1986] Brooks, R.A. (1986). Acheiving artificial intelligence through building robots. A.I.Memo 899, M.I.T.

[Brooks, 1989] Brooks, R.A. (February 1989). A robot that walks; emergent behaviours from a carefully evolved network. A.I. Memo 1091, MIT.

[Brooks, 1991a] Brooks, R.A. (1991a). Challenges for complete creature architectures. In Meyer, J-A. and Wilson, S.W., (eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, MIT Press.

[Brooks, 1991b] Brooks, R.A. (1991b). Elephants don't play chess. In Maes, P., (ed.), *Designing Autonomous Agents.* MIT Press, Cambridge, Mass.

[Brooks, 1991c] Brooks, R.A. (1991c). Intelligence without reason. In *Proceedings of IJCAI-91*, Morgan Kaufmann.

[Brooks, 1991d] Brooks, R.A. (1991d). Intelligence without representation. *Artificial Intelligence*, 47:139–159.

[Brooks, 1992] Brooks, R.A. (1992). Artificial life and real robots. In *Proceedings of the First European Conference on Artificial Life*. MIT Press.

[Camhi, 1987] Camhi, J.M. (1987). How is a toad not like a bug? *The Behavioural and Brain Sciences*, 10:371.

[Camhi, 1988] Camhi, J.M. (1988). Inverterbrate neuroethology. *Experientia*, 44:361–362.

[Cecconi & Parisi, 1991] Cecconi, F. and Parisi, D. (1991). Evolving organisms that can reach for objects. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour*. MIT Press.

[Chan & Tidwell, 1992] Chan, K.H. and Tidwell, P.M. (1992). The reality of artificial life: Can computer simulations become realizations. In *Proceedings of the Third Artificial Life Conference*.

[Clancey, 1991] Clancey, W.J. (1991). Situated cognition: Stepping out of representational flatland. *AICOM*, 4:109–112.

[Clancey, 1992] Clancey, W.J. (1992). The frame of reference problem in the design of intelligent machines. In von Lehn, K., (ed.), *Architectures for Intelligence*. Erlbaum, Hillsdale, NJ.

[Clark, 1989] Clark, A. (1989). *Microcognition*. M.I.T. Press, Cambridge MA.

[Cliff, 1990] Cliff, D.T. (1990). Computational neuroethology: A provisional manifesto. CSRP 162, University of Sussex.

[Cliff, 1992] Cliff, D.T. (1992). Neural networks for visual tracking in an artificial fly. In *Proceedings of the First European Conference on Artificial Life.* MIT Press.

[Cliff *et al*, 1992] Cliff, D.T., Husbands, P. and Harvey, I. (1992). Evolving visually guided robots. CSRP 220, University of Sussex.

[Coderre, 1989] Coderre, B. (1989). Modeling behaviour in petworld. In Langton, C.G., (ed.), *Artificial Life*, Redwood City, California. Addison-Wesley.

[Cohen, 1991] Cohen, P.R. (1991). A survey of the Eighth National Conference on Artificial Intelligence: Pulling together or pulling apart. *AI Magazine*, pages 17–41.

[Colles, 1992] Colles, T. (1992). Using the PSI Mini Module Running ABOS. Unpublished technical document.

[Connell, 1987] Connell, Jonathan H. (1987). Creature design with the subsumption architecture. *IJCAI-87*, pages 1124–1126.

[Crick, 1989] Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337:129–132.

[Cruse, 1991] Cruse, H. (1991). Coordination of leg movement in walking animals. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour.* MIT Press.

[Cummins, 1989] Cummins, R. (1989). *Meaning and Mental Representation.* MIT Press, Cambridge, MA.

[Daily *et al*, 1988] Daily, M., Harris, J., Keirsey, D., Olin, K., Payton, D. W., Reiser, K., Rosenblatt, J., Tseng, D. and Wong, V. (1988). Autonomous cross-country navigation with the ALV. In *Proccedings of IEEE International Conference on Advanced Robotics.*

[Dallas, 1990] Dallas, Jim. (1990). Co-operative search behaviour in a group of lego vehicles. Unpublished M.Sc. thesis, Dept. of Artificial Intelligence, University of Edinburgh.

[de Garis, 1992] de Garis, H. (1992). Steerable GenNets. In *Proceedings of the First European Conference on Artificial Life.* MIT Press.

[Dennett, 1979] Dennett, D.C. (1979). *Brainstorms:Philosophical Essays on Mind and Psychology.* The Harvester Press, Sussex.

[Di Stefano *et al*, 1990] Di Stefano, G., Iacino, L. and Schiff, H. (1990). What the Mantis shrimp's eye (possibly) tells its raptorial appendages. *Biological Cybernetics*, 63:333–401.

[Doherty, 1985] Doherty, J.A. (1985). Trade-off phenomena in calling song recognition and phonotaxis in the cricket, *Gryllus bimaculatus (Orthoptera Gryllidae). Journal of Comparative Physiology A*, 156:787–801.

[Doherty, 1991] Doherty, J.A. (1991). Song recognition and localization in the phonotaxis behaviour of the field cricket, *Gryllus bimaculatus (Orthoptera Gryllidae). Journal of Comparative Physiology A*, 168:213–222.

[Donnett, 1993] Donnett, J. (1993). Unpublished Ph.D. thesis, University of Edinburgh.

[Donnett & Smithers, 1990] Donnett, J. and Smithers, T. (1990). Lego vehicles: A technology for studying intelligent systems. Research Paper 490, University of Edinburgh.

[Eisenstein & Reep, 1985] Eisenstein, E.M. and Reep, R.L. (1985). Learning. In Kerkut, G.A. and Gilbert, L.I., (eds.), *Comprehensive Insect Physiology, Biochemistry and Pharmacology, volume 9: Behaviour.* Pergamon Press, Oxford.

[Erulkar, 1972] Erulkar, S.D. (1972). Comparative aspects of spatial localization of sound. *Physiological Reviews*, 52:237–360.

[Esch *et al*, 1980] Esch, H., Huber, F. and Wohlers, D. (1980). Primary auditory interneurons in crickets: Physiology and central projections. *Journal of Comparative Physiology A*, 137:27–38.

[Ewert, 1987] Ewert, J.P. (1987). Neuroethology of releasing mechanisms: Prey catching in toads. *The Behavioural and Brain Sciences*, 10:337–405.

[Ewert *et al*, 1981] Ewert, J.P., Capranica, R.R. and Ingle, D.J., (eds.). (1983). *Advances in Vertebrate Neuroethology.* Plenum, New York.

[Ewing, 1989] Ewing, A.W. (1989). *Arthropod Bioacoustics: Neurobiology and Behaviour.* Edinburgh Universtity Press, Edinburgh.

[Fernald, 1984] Fernald, R.D. (1984). Neuroethology according to hoyle. *The Behavioural and Brain Sciences*, 7:388.

[Ferretti & de Talens, 1975] Ferretti, C. Taddei and de Talens, A. Fernandez Perez. (1975). Landing and optomotor responses of the fly Musca. In Horridge, G.A., (ed.), *The compound eye and vision of insects.* Clarendon Press, Oxford.

[Fletcher & Thwaites, 1979] Fletcher, N.H. and Thwaites, S. (1979). Acoustical analysis of the auditory system of the cricket *Teleogryllus commodus (Walker). Journal of the Acoustical Society of America*, 66:350–357.

[Flynn & Brooks, 1989] Flynn, A. M. and Brooks, R.A. (1989). Battling reality. A.I.Memo 1148, M.I.T.

[Fodor, 1981] Fodor, J.A. (1981). Methodological solipsism. In Haugeland, J., (ed.), *Mind design.* MIT Press, Cambridge, Mass.

[Franceschini *et al*, 1989] Franceschini, N., Riehle, A. and LeNestour, A. (1989). Directionally selective motion detection by insect neurons. In Stavenga and Hardie, (eds.), *Facets of Vision.* Springer-Verlag, Berlin.

[Franceschini *et al*, 1991] Franceschini, N., Pichon, J-M. and Blanes, C. (1991). Real time visuomotor control: from flies to robots. In *Proccedings of IEEE Fifth International Conference on Advanced Robotics.*

[French & Cade, 1987] French, B.W. and Cade, W.H. (1987). The timing of calling, movement and mating in the field crickets *Gryllus veletis, Gryllus pennsylvanicus and Gryllus integer. Behavioural Ecology and Sociobiology*, 21:157–162.

[Geman *et al*, 1992] Geman, S., Bienenstock, E. and Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, 4:1–58.

[Gibson, 1961] Gibson, J.J. (1961). Ecological optics. *Vision Research*, 1:253–262.

[Gibson, 1966] Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems.* Houghton Mifflin, Boston.

[Gibson, 1979] Gibson, J.J. (1979). *The ecological approach to visual perception.* Houghton Mifflin, Boston.

[Hansell, 1985] Hansell, M.H. (1985). Ethology. In Kerkut, G.A. and Gilbert, L.I., (eds.), *Comprehensive Insect Physiology, Biochemistry and Pharmacology, volume 9: Behaviour.* Pergamon Press, Oxford.

[Harnad, 1990] Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

[Heiligenburg, 1987] Heiligenburg, W. (1987). The control of behavioral performance in networks processing temporal and spatial patterns of sensory information. In Guthrie, D.M., (ed.), *Aims and Methods in Neuroethology.* Manchester University Press, Manchester.

[Henderson & Grupen, 1990] Henderson, T.C. and Grupen, R. (1990). Logical behaviours. *Journal of Robotic Systems*, 7:309–336.

[Hennig, 1988] Hennig, R.M. (1988). Ascending auditory interneurons in the cricket *Teleogryllus commodus (Walker)*: comparative physiology and direct connections with afferents. *Journal of Comparative Physiology A*, 163:135–143.

[Hill & Boyan, 1977] Hill, K.G. and Boyan, S. (1977). Sensitivity to frequency and direction of sound in the auditory system of crickets (*Gryllidae*). *Journal of Comparative Physiology A*, 121:79–97.

[Hissman, 1990] Hissman, K. (1990). Strategies of mate finding in the European field cricket (*Gryllus Campestris*) at different population densities: a field study. *Ecological Entomology*, 15:281–291.

[Horn, 1992] Horn, J. (1992). Measuring the evolving complexity of stimulus-response organisms. In *Proceedings of the First European Conference on Artificial Life*. MIT Press.

[Horswill & Brooks, 1988] Horswill, I.D. and Brooks, R. (1988). Situated vison in a dynamic world: Chasing objects. In *Proceedings of AAAI-88*, MIT Press.

[Horswill, 1992] Horswill, I. (1992). Characterizing adaptation by constraint. In *Proceedings of the First European Conference on Artificial Life*, MIT Press.

[Hoyle, 1983] Hoyle, G. (1983). On the way to neuroethology: The identified neuron approach. In Huber, F. and Markl, H., (eds.), *Neuroethology and Behavioural Physiology*. Springer-Verlag, Berlin.

[Hoyle, 1984] Hoyle, G. (1984). The scope of neuroethology. *The Behavioural and Brain Sciences*, 7:367–412.

[Huber, 1983] Huber, F. (1983). Neural correlates of orthopteran and cicada phonotaxis. In Huber, F. and Markl, H., (eds.), *Neuroethology and Behavioural Physiology*. Springer-Verlag, Berlin.

[Huber, 1988] Huber, F. (1988). Invertebrate neuroethology—guiding principles. *Experientia*, 44:428–431.

[Huber & Thorson, 1985] Huber, F. and Thorson, J. (1985). Cricket auditory communication. *Scientific American*, 253:6:47–54.

[Huber *et al*, 1984] Huber, F., Kleindienst, H.U., Weber, T. and Thorson, J. (1984). Auditory behaviour of the cricket III Tracking of male calling song by surgically and developmentally one-eared females, and the curious role of the anterior tympanum. *Journal of Comparative Physiology A*, 155:725–738.

[Hurlbert & Poggio, 1988] Hurlbert, A. and Poggio, T. (1988). Making machines and Artificial Intelligence see. In Graubard, S.R., (ed.), *The Artificial Intelligence debate: False Starts, Real Foundations.* M.I.T. Press, Cambridge MA.

[Iacino *et al*, 1990] Iacino, L., Di Stefano, G. and Schiff, H. (1990). A neural model for localizing targets in space accomplished by the eye of a mantis shrimp. *Biological Cybernetics*, 63:383–391.

[Jamon, 1991] Jamon, M. (1991). The contribution of quantitative models to the long distance orientation problems. In Meyer, J-A. and Wilson, S.W., (eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, MIT Press.

[Kaebling, 1992] Kaebling, L.P. (1992). An adaptable mobile robot. In *Proceedings of the First European Conference on Artificial Life*, MIT Press.

[Kay, 1992] Kay, J.W., (1992). Artificial neural networks and statistics. Talk given at Edinburgh University.

[Kleindienst *et al*, 1981] Kleindienst, H.U., Koch, U.T. and Wohlers, D.W. (1981). Analysis of the cricket auditory system by acoustic stimulation using a closed sound field. *Journal of Comparative Physiology A*, 141:283–296.

[Koch & Segev, 1989] Koch, C. and Segev, I., (eds.). (1989). *Methods in Neuronal Modeling.* M.I.T. Press - Bradford Books, Cambridge MA.

[Koditschek & Buhler, 1991] Koditschek, D.E. and Buhler, M. (1991). Analysis of a simplified hopping robot. *The International Journal of Robotics Research*, 10:587–605.

[Konishi, 1993] Konishi, M. (1993). Listening with two ears. *Scientific American*, April:34–41.

[Koza, 1992] Koza, J.R. (1992). Evolution of subsumption using genetic programming. In *Proceedings of the First European Conference on Artificial Life*. MIT Press.

[Kramer, 1976] Kramer, E. (1976). The orientation of walking honeybees in odour fields with small concentration gradients. *Physiological Entomology*, 1:27–37.

[Krasne & Wine, 1987] Krasne, F.B. and Wine, J.J. (1987). Evasion responses of the crayfish. In Guthrie, D.M., (ed.), *in Aims and Methods in Neuroethology*. Manchester University Press, Manchester.

[Kreigman *et al*, 1987] Kreigman, D.J., Triedl, E. and Binford, T.O. (1987). A mobile robot: sensing, planning and locomotion. In *Proccedings of IEEE International Conference on Robotics and Automation*.

[Kupferman, 1984] Kupferman, I. (1984). They really are complex when you get to know them. *The Behavioural and Brain Sciences*, 7:393.

[Lakoff, 1988] Lakoff, G. (1988). Smolensky, semantics and the sensorimotor system. *The Behavioural and Brain Sciences*, 11:39-40.

[Langton, 1989] Langton, C.G., (ed.). (1989). *Artificial Life*, Redwood City, California. Addison-Wesley.

[Larsen & Michelson, 1978] Larsen, O.N. and Michelson, A. (1978). Biophysics of the ensiferan ear III The cricket ear as a four-input system. *Journal of Comparative Physiology*, 123:217–227.

[Larsen *et al*, 1983] Larsen, O.N., Surlykke, A. and Michelson, A. (1983). Directionality of the cricket ear: A property of the tympanal membrane. *Journal of Comparative Physiology*, 153.

[Lee, 1980] Lee, D.N. (1980). The optic flow field: the foundation of vision. *Philosophical Transactions of the Royal Society of London B*, 290:169–179.

[Lesperance & Levesque, 1990] Lesperance, Y. and Levesque, H.J. (1990). Indexical knowledge in robot plans. In *Proceedings of AAAI-90.* MIT Press.

[Maes & Brooks, 1990] Maes, Pattie and Brooks, Rodney A. (1990). Learning to coordinate behaviors. In *Proceedings of AAAI -90*, MIT Press.

[Malcolm, 1990] Malcolm, C. (1990). Behavioural modules in robotic assembly. Edinburgh University, Dept. of AI (unpublished).

[Malcolm *et al*, 1989] Malcolm, C., Smithers, T. and Hallam, J. (1989). An emerging paradigm in robot architecture. DAIRP 447, Edinburgh University.

[Mardia, 1972] Mardia, K.V. (1972). *Statistics of Directional Data.* Academic Press, London.

[Maturana & Varela, 1980] Maturana, H.R. and Varela, F.J. (1980). *Autopoesis and cognition.* Reidel, Dordrecht.

[Maze, 1981] Maze, J. (1981). *The Meaning of Behaviour.* Allen and Unwin, London.

[McFarland, 1991] McFarland, D. (1991). What it means for robot behaviour to be adaptive. In Meyer, J-A. and Wilson, S.W., (eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, MIT Press.

[Mein, 1991] Mein, R. (1991). Cooperative behaviour in uniformly and differentially programmed Lego vehicles. Unpublished M.Sc. thesis, Department of Artificial Intelligence.

[Mendenhall *et al*, 1981] Mendenhall, W., Scheaffer, R.L. and Wackerly, D.D. (1981). *Mathematical Statistics with Applications.* Duxberry Press, Boston, Mass.

[Meyer & Guillot, 1991] Meyer, J.A. and Guillot, A. (1991). Simulation of adaptive behaviour in animats: Review and prospect. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour.* MIT Press.

[Meyer & Wilson, 1991] Meyer, J.A. and Wilson, S.W., (eds.). (1991). *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour.* MIT Press.

[Michel, 1986] Michel, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin*, 100:398–407.

[Michelson & Larsen, 1985] Michelson, A. and Larsen, O. (1985). Hearing and sound. In Kerkut, G.A. and Gilbert, L.I., (eds.), *Comprehensive Insect Physiology, Biochemistry and Pharmacology.* Pergammon Press, Oxford.

[Michelson, 1983] Michelson, A. (1983). Biophysical bases of sound communication. In Lewis, B., (ed.), *Bioacoustics - A Comparative Approach.* Academic Press, London.

[Miller, 1983] Miller, L.A. (1983). How insects detect and avoid bats. In Huber, F. and Markl, H., (eds.), *Neuroethology and Behavioural Physiology.* Springer-Verlag, Berlin.

[Moravec, 1983] Moravec, H.P. (1983). The Stanford Cart and the CMU Rover. *Proceedings of the IEEE*, 71:872–884.

[Moravec, 1984] Moravec, Hans. (1984). Locomotion, vision and intelligence. In *Robotics Research.* MIT Press.

[Mörchen, 1980] Mörchen, A. (1980). Spike count and response latency. Two basic parameters encoding sound direction in the CNS of insects. *Naturwissenschaft*, 65:656-657.

[Naylor *et al*, 1966] Moravec, Hans. (1966). Computer Simulation Techniques. John Wiley and Sons.

[Nehmzow & Smithers, 1992] Nehmzow, Ulrich and Smithers, Tim. (1992). Learning multiple competences: some initial experiments. In *Proceedings of ECAI-92.*

[Newell & Simon, 1976] Newell, A. and Simon, H.A. (1976). Computer science as an empirical enquiry: Symbols and search. *Communications of the ACM*, 19:113–126.

[Nilsson, 1984] Nilsson, N.J. (1984). Shakey the robot. A.I. Center Technical Note 323, SRI.

[Oldfield *et al*, 1986] Oldfield, B.P., Kleindienst, H.U. and Huber, F. (1986). Physiology and tonotopic organization of auditory receptors in the cricket *Gryllus bimaculatus De Geer. Journal of Comparative Physiology A*, 152:457–464.

[Payton, 1986] Payton, D. W. (1986). An architecture for reflexive autonomous vehicle control. In *Proccedings of IEEE International Conference on Advanced Robotics.*

[Payton, 1991] Payton, David W. (1991). Internalized plans: A representation for action resources. In Maes, Pattie, (ed.), *Designing Autonomous Agents*, MIT Press.

[Pebody, 1991] Pebody, Miles. (1991). How to make a Lego robot do the right thing. Unpublished M.Sc. thesis, Department of Artificial Intelligence.

[Pfeifer and Verschure, 1992] Pfeifer, R. and Verschure, P. (1992). The challenge of autonomous agents: pitfalls and how to avoid them. Prepared from the

Workshop on Emergence, Situatedness, Subsumption and Symbol Grounding at Corsendonk Priory, Belgium.

[Pichon *et al*, 1989] Pichon, J-M., Blanes, C. and Francescini, N. (1989). Visual guidance of a mobile robot equipped with a network of self motion sensors. *SPIE*, 1195 Mobile Robots IV:44–53.

[Pierce & Kuipers, 1991] Pierce, D. and Kuipers, B. (1991). Learning hill-climbing functions as a strategy for generating behaviours in a mobile robot. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour*. MIT Press.

[Pierce & Ollanson, 1987] Pierce, G.J. and Ollanson, J.G. (1987). Eight reasons why optimal foraging theory is a complete waste of time. *Oikos*, 49:111–125.

[Pollack, 1986] Pollack, G.S. (1986). Discrimination of calling song models by the cricket, *Teleogryllus oceanicus*: the influence of sound direction on neural encoding of the stimulus temporal pattern and on phonotactic behaviour. *Journal of Comparative Physiology A*, 158:549–561.

[Pollack & Hoy, 1981] Pollack, G.S. and Hoy, R. (1981). Phonotaxis to individual rhythmic components of a complex cricket calling song. *Journal of Comparative Physiology*, 144:367–373.

[Pollack & Plourde, 1982] Pollack, G.S. and Plourde, N. (1982). Directionality of acoustic orientation in flying crickets. *Journal of Comparative Physiology A*, 146:207–215.

[Popov & Shuvalov, 1977] Popov, A.V. and Shuvalov, V.F. (1977). Phonotactic behaviour of crickets. *Journal of Comparative Physiology*, 119:111–126.

[Powers, 1973] Powers, W.T. (1973). *Behaviour: the Control of Perception*. Aldine, Chicago.

[Pylyshyn, 1987] Pylyshyn, Z.W., (ed.). (1987). *The Robot's dilemma: the frame problem in Artificial Intelligence*. Ablex, Norwood, NJ.

[Raibert, 1986] Raibert, M.H. (1986). Running with symmetry. *The International Journal of Robotics Research*, 5:3–19.

[Raibert, 1989] Raibert, M.H. (1989). Dynamically stable legged locomotion. AI-TR 1179, MIT.

[Rasnow *et al*, 1988] Rasnow, B., Assad, C., Nelson, M.E. and Bower, J.M. (1988). Simulation and measurement of the electric fields generated by weakly electric fish. In Touretzky, D.S., (ed.), *Advances in Neural Information Processing Systems*. Morgan Kaufman, Palo Alto, CA.

[Reeke & Edelman, 1988] Reeke, G.N. and Edelman, G.M. (1988). Real brains and Artificial Intelligence. In Graubard, S.R., (ed.), *The Artificial Intelligence debate: False Starts, Real Foundations*. M.I.T. Press, Cambridge MA.

[Regen, 1913] Regen, J. (1913). Uber die Anlockung des Weibchens van Gryllus campestris L. durch telephonisch ubertragene Stridulationslaute des Manchens. *Pflugers Arch.*, 155:193–200.

[Resnick, 1991] Resnick, M. (1991). *Logo: Animal simulations and the restructuring of science education. In Meyer, J-A. and Wilson, S.W., (eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, MIT Press.

[Roberts, 1987] Roberts, A. (1987). Skin sensory modalities, free nerve endings and behaviour: a reappraisal based on studies of amphibian embryos. In Guthrie, D.M., (ed.), *Aims and Methods in Neuroethology*. Manchester University Press, Manchester.

[Romer & Rheinlaender, 1983] Romer, H. and Rheinlaender, J. (1983). Electrical stimulation of the tympanal nerve as a tool for analysing the responses of auditory interneurons in the cricket. *Journal of Comparative Physiology A*, 152:289–296.

[Romer, 1985] Romer, H. (1985). Anatomical representation of frequency and intensity in the auditory system of *Orthoptera*. In Kalmring, K. and Elsner, N., (eds.), *Acoustic and Vibrational Communication in Insects*. Verlag Paul Parey, Berlin.

[Rosen, 1983] Rosen, R. (1983). Role of similarity principles in data extrapolation. *American Journal of Physiology*, 13:591–599.

[Rothenberg, 1989] Rothenberg, J. (1989). The nature of modeling. In Widman, L.E., Loparo, K.A. and Nelson, N.R., (eds.), *Artificial Intelligence, Simulation and Modeling*. John Wiley and Sons, New York.

[Ryle, 1949] Ryle, G. (1949). *The Concept of Mind*. Brendon and Son, London.

[Schildberger, 1984] Schildberger, K. (1984). Temporal selectivity of identified auditory neurons in the cricket brain. *Journal of Comparative Physiology A*, 155:171–185.

[Schildberger, 1985] Schildberger, K. (1985). Recognition of temporal patterns by identified auditory neurons in the cricket brain. In Kalmring, K. and Elsner, N., (eds.), *Acoustic and Vibrational Communication in Insects*. Verlag Paul Parey, Berlin.

[Schildberger, 1988] Schildberger, K. (1988). Behavioural and neuronal methods of cricket phonotaxis. *Experientia*, 44:408–415.

[Schildberger & Horner, 1988] Schildberger, K. and Horner, M. (1988). The function of auditory neurons in cricket phonotaxis I Influence of hyperpolarization of identified neurons on sound localization. *Journal of Comparative Physiology A*, 163:621–631.

[Schmitz, 1985] Schmitz, B. (1985). Phonotaxis in *Gryllus campestris L.* III Intensity dependence of the behavioural performance and relative importance of tympana and spiracles in directional hearing. *Journal of Comparative Physiology A*, 156:165–180.

[Schmitz *et al*, 1982] Schmitz, B., Scharstein, H. and Wendler, G. (1982). Phonotaxis in *Gryllus campestris L.* I Mechanism of acoustic orientation in intact female crickets. *Journal of Comparative Physiology A*, 148:431–444.

[Schmitz *et al*, 1988] Schmitz, B., Kleindienst, H.U. and Huber, F. (1988). Acoustic orientation in adult female crickets after unilateral foreleg amputation in the larva. *Journal of Comparative Physiology A*, 162:715–728.

[Schwartz, 1988] Schwartz, J.T. (1988). The new connectionism: developing relationships between neuroscience and Artificial Intelligence. In Graubard, S.R., (ed.), *The Artificial Intelligence debate: False Starts, Real Foundations.* M.I.T. Press, Cambridge MA.

[Simmons, 1988] Simmons, L.W. (1988). The calling song of the field cricket, *Gryllus bimaculatus (De Geer)*: constraints on transmission and its role in intermale competition and female choice. *Animal Behaviour*, 36:380–394.

[Singh *et al*, 1991] Singh, S., Feng, D., Keller, P., Schaffer, G., Schi, W.F., Shin, D.H, West, J. and Wu, B.X. (1991). A system for fast navigation of autonomous vehicles. RI-TR 91-20, CMU.

[Skinner, 1977] Skinner, B.F. (1977). Why I am not a cognitive psychologist. *Behaviourism*, 5:1–12.

[Skinner, 1984] Skinner, B.F. (1984). Behaviourism at fifty. *(reprinted in) Behavioural and Brain Sciences*, 7:615–667.

[Smith, 1985] Smith, Brian. (1985). Prologue to the reflection and semantics in a procedural language. In *Readings in Knowledge Representation.* Morgan Kaufman.

[Smithers, 1992] Smithers, T. (1992). Are autonomous agents information processing systems? Prepared from the Workshop on Emergence, Situatedness, Subsumption and Symbol Grounding at Corsendonk Priory, Belgium.

[Stabel *et al*, 1989] Stabel, J., Wendler, G. and Scharstein, H. (1989). Cricket phonotaxis — localization depends on recognition of the calling song pattern. *Journal of Comparative Physiology A*, 165:165–177.

[Stout & McGhee, 1988] Stout, J.F. and McGhee, R. (1988). Attractiveness of the male *Acheta domestica* calling song to females II The relative importance of syllable period, intensity and chirp rate. *Journal of Comparative Physiology A*, 164:277–287.

[Stout *et al*, 1983] Stout, J.F., DeHann, C.H. and McGhee, R. (1983). Attractiveness of the male *Acheta domestica* calling song to females. *Journal of Comparative Physiology A*, 153:509–521.

[Stout *et al*, 1985] Stout, J.F, Atkins, G. and Burghardt, F. (1985). The characterization and possible importance for phonotaxis of L-shaped ascending acoustic interneurons in the cricket (*Acheta Domesticus*). In Kalmring, K. and Elsner, N., (eds.), *Acoustic and Vibrational Communication in Insects*. Verlag Paul Parey, Berlin.

[Stout *et al*, 1987] Stout, J.F, Atkins, G., Weber, T. and Huber, F. (1987). The effect of visual input on calling song attractiveness for female *Acheta domesticus*. *Physiological Entomology*, 12:135–140.

[Stout *et al*, 1988] Stout, J.F., DeHaan, C.H., Hall, J.C. and Rhodes, M. (1988). Processing of calling songs by an L-shaped neuron in the prothoracic ganglion of the female cricket, *Acheta domesticus*. *Physiological Entomology*, 13:89–101.

[Thompson *et al*, 1992] Thompson, E., Palacios, A. and Varela, F.J. (1992). Ways of coloring: Comparative color vision as a case study for cognitive science. *Behavioural and Brain Sciences*, 15:1–74.

[Thorson *et al*, 1982] Thorson, J., Weber, T. and Huber, F. (1982). Auditory behaviour in the cricket II Simplicity of calling song recognition in *Gryllus* and anomolous phonotaxis at abnormal carrier frequencies. *Journal of Comparative Physiology A*, 146:361–378.

[Tyrrell & Mayhew, 1991] Tyrrell, T. and Mayhew, J.E.W. (1991). Computer simulation of an animal environment. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour.* MIT Press.

[Ullman, 1980] Ullman, S. (1980). Against direct perception. *Behavioural and Brain Sciences*, 3:373–415.

[Varela & Bourgine, 1992] Varela, F.J. and Bourgine, P., (eds.). (1992). *Towards a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life.* MIT Press.

[Varela *et al*, 1991] Varela, F.J., Thompson, E. and Palacios, A. (1991). *The embodied mind.* MIT Press, Boston, Mass.

[vonHelverson & vonHelverson, 1983] von Helverson, D. and von Helverson, O. (1983). Species recognition and acoustic localization in acridid grasshoppers: a behavioural approach. In Huber, F. and Markl, H., (eds.), *Neuroethology and Behavioural Physiology.* Springer-Verlag, Berlin.

[vonHelverson & vonHelverson, 1987] von Helverson, O. and von Helverson, D. (1987). Innate reciever mechanisms in the acoustic communication of orthopteran insects. In Guthrie, D.M., (ed.), *in Aims and Methods in Neuroethology.* Manchester University Press, Manchester.

[Walter, 1953] Walter, W.G. (1953). *The Living Brain.* Duckworth.

[Watson, 1913] Watson, J.B. (1913). Psychology as the behaviourist views it. *Psychological Review*, 20:158–177.

[Webb, 1987] Webb, B.H. (1987). The nature of cognitive simulation. Unpublished Honours Thesis, University of Sydney.

[Weber & Thorson, 1988] Weber, T. and Thorson, J. (1988). Auditory behaviour in the cricket II Interaction of direction of tracking with perceived temporal

pattern in split-song paradigms. *Journal of Comparative Physiology A*, 163:13–22.

[Weber, 1984] Weber, T. (1984). Acoustical recognition in crickets. In Varju, D. and Schnitzler, H.U., (eds.), *Localization and orientation in biology and engineering.* Springer, Berlin.

[Weber *et al*, 1981] Weber, T., Thorson, J. and Huber, F. (1981). Auditory behaviour in the cricket I Dynamics of compensated walking and discrimination paradigms on the Kramer treadmill. *Journal of Comparative Physiology*, 141:215–232.

[Wehner, 1987] Wehner, R. (1987). Matched filters—neural models of the external world. *Journal of Comparative Physiology A*, 161:511–531.

[Whittaker *et al*, 1991] Whittaker, W., Kanade, T. and Mitchell, T. (1991). 1990 Year End Report: Autonomous Planetary Rover at Carnegie Mellon. RI-TR 91-19, CMU.

[Wiese & Eilts-Grimm, 1985] Wiese, K. and Eilts-Grimm, K. (1985). Functional potential of recurrent lateral inhibition in cricket audition. In Kalmring, K. and Elsner, N., (eds.), *Acoustic and Vibrational Communication in Insects.* Verlag Paul Parey, Berlin.

[Wilson, 1966] Wilson, D.M. (1966). Insect walking. *Annual Review of Entomology*, 11:103–122.

[Wilson, 1985] Wilson, S.W. (1985). Knowledge growth in an artificial animal. In *Proceedings of the International Conference on genetic Algorithms and Their Applications.* Carnegie Mellon University.

[Wilson, 1991] Wilson, Stewart W. (1991). The Animat path to AI. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour.* MIT Press.

[Winograd & Flores, 1986] Winograd, T. and Flores, F. (1986). *Understanding computers and cognition: A new foundation for design.* Ablex Publishing Co., Norwood NJ.

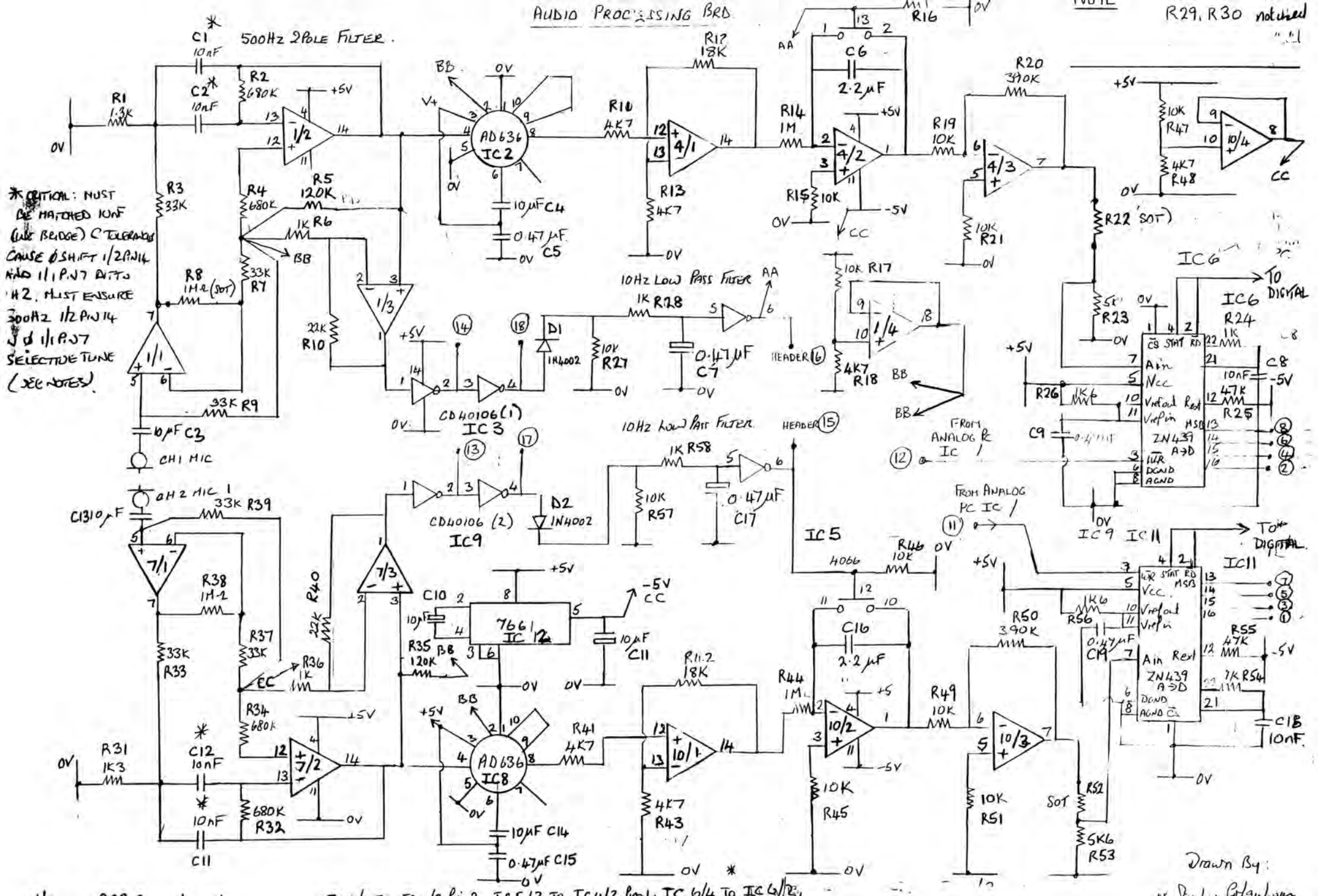
[Wohlers & Huber, 1982] Wohlers, D.W. and Huber, F. (1982). Processing of sound signals by six types of neurons in the prothoracic ganglion of the cricket, *Grullus campestris L. Journal of Comparative Physiology*, 146:161–173.

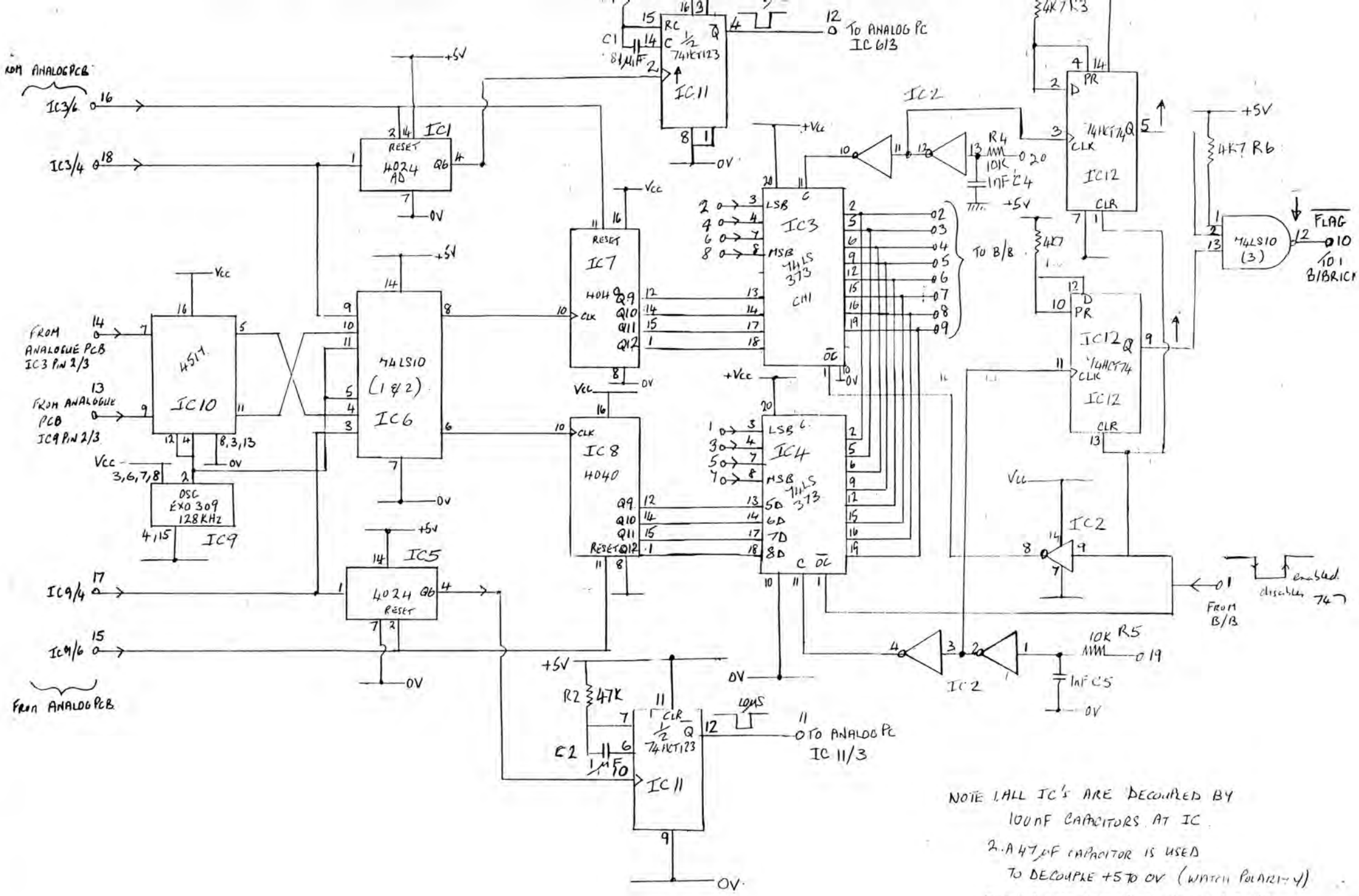
[Zuk, 1987] Zuk, M. (1987). Variability in attractiveness of male field crickets (*Orthoptera: Gryllidae*) to females. *Animal Behaviour*, 35:1220–1248.

# Appendix A

# Diagrams of Auditory Circuit Boards

Drawn by Sandy Colquhoun


AUDIO PROCESSING BRD
C1 10nF
C2 10nF
500Hz 2POLE FILTER.
R1 1.3K
R2 680K
R3 33K
R4 680K
R5 120K
1K R6
R7 33K
R8 1M2 (SOT)
R9 33K
10µF C3
CH1 MIC
0V
+5V
-5V
1/1
1/2
1/3
AD636 IC2
10µF C4
0.47µF C5
R10 4K7
R11 18K
R12 18K
R13 4K7
4/1
4/2
4/3
R14 1M
R15 10K
R16
C6 2.2µF
R17 10K
R18 4K7
R19 10K
R20 390K
R21 10K
R22 (SOT)
R23 5K
10Hz LOW PASS FILTER
1K R28
0.47µF C7
D1 IN4002
R27 10K
CD40106(1) IC3
1/4
10/4
IC6
TO DIGITAL
R24 1K
R25 47K
R26 1K
C8 10nF
C9 0.47µF
ZN439 A→D
Ain Vcc Vrefout Vrefin Rest MSB LSB DGND AGND
IC9
IC11
R46 10K
R47 10K
R48 4K7
CC
HEADER (6)
HEADER (15)
FROM ANALOG PC IC
TO DIGITAL
BB
AA
R29, R30 notched
* CRITICAL: MUST BE MATCHED 10nF (USE BRIDGE) C TOLERANCES CAUSE ØSHIFT 1/2 PIN 14 AND 1/1 PIN 7 DITTO H2. MUST ENSURE 500Hz 1/2 PIN 14 AND 1/1 PIN 7 SELECTIVE TUNE (SEE NOTES).
aH2 MIC 1
C13 10µF
33K R39
7/1
7/2
7/3
R38 1M2
R37 33K
R33 33K
EC
R36 1K
R34 680K
R40 22K
R31 1K3
C12 10nF
10nF
680K R32
C11
CD40106 (2) IC9
D2 IN4002
R35 120K
7661 IC12
C10 10nF
10µF C11
AD636 IC8
10µF C14
0.47µF C15
R41 4K7
R42 18K
R43 4K7
10/1
10/2
10/3
R44 1M
R45 10K
C16 2.2µF
R49 10K
R50 390K
R51 10K
R52 SOT
5K6 R53
IC5
4066
R46 10K
0.47µF C17
1K R58
10K R57
0.47µF C18
R55 47K
7K R54
R56 1K6
0.47µF C19
Drawn By:


FROM ANALOGUE PCB
IC3/6
IC3/4
IC1
4024 AD
RESET
Q6
+5V
0V
C1
8.2µF
IC11
74HCT123
RC
C
Q
1/2
TO ANALOG PC IC 613
IC7
4014
Q9
Q10
Q11
Q12
CLK
Vcc
IC3
LSB
MSB
74LS 373
CMI
OC
TO B/B
IC2
R4
10K
1nF C4
4K7 R3
4K7 R6
74HCT74
PR
CLR
IC12
74LS10 (3)
FLAG
TO BIBRICK
FROM ANALOGUE PCB IC3 PIN 2/3
FROM ANALOGUE PCB IC9 PIN 2/3
4511
IC10
74LS10 (1&2)
IC6
OSC EXO 309 128KHz
IC9
IC8
4040
IC4
IC5
4024
IC9/4
IC9/6
FROM ANALOG PCB
R2 47K
C2 1µF
IC11
TO ANALOG PC IC 11/3
10K R5
1nF C5
FROM B/B
NOTE 1. ALL IC's ARE DECOUPLED BY 100nF CAPACITORS AT IC
2. A 47µF CAPACITOR IS USED TO DECOUPLE +5 TO 0V (WATCH POLARITY)
3. ONLY ONE 4K7 PULL USED ON IC 12

# Appendix B

## B.1 Code for robot interface

```
/* functions for robot interface */
extern int l_amp, r_amp, l_phase, r_phase;
extern int store[2500][6], index;
extern byte motorstate;

/*-----------------------------------------------------------------*/

/* INIT initialises the variables used in the following functions */

init()
{
freeze; /*motors are off*/
stop_counters;
high_access;
start_counters;
}

/*-----------------------------------------------------------------*/

/* SOUND */

sound()
{
byte lbyte, rbyte;

stop_counters; /*sufficient time for reading should have elapsed
since last function call*/

if flag       /*  wasn't enough sound for a reading */
l_phase = r_phase = l_amp = r_amp = 0;
else       /* read data from ears */
{
lbyte = ears;
low_access;
rbyte = ears;
high_access; /* reset channels and flag */
l_phase = (lbyte>>4);
l_amp = (lbyte&0x0F);
```

```
r_phase = (rbyte>>4);
r_amp = (rbyte&0x0F);
}
if (index<2500)
        {
        store[index][1] = l_phase;
        store[index][2] = l_amp;
        store[index][3] = r_phase;
        store[index][4] = r_amp;
        store[index][0] = (motorstate>>7);
        store[index][5] = (int)(gettime()>>2);
        index += 1;
        }
lights(motorstate);
start_counters; /* begin next reading */
}



/*-------------------------------------------------------------*/

/* MOVE transfers movement commands to the motors. It should only
be called on when the movement changes */

move(manner)
{
if (manner==again)
        {
make_move;
        }
if (manner==forward)
        {
        leftfor;
        rightfor;
        }
if (manner==turnleft)
{
        leftstop;
        }
if (manner==not)
        {
freeze;
        }
if (manner==turnright)
        {
        rightstop;
        }
if (manner==back)
        {
        leftrev;
        rightrev;
        }
```

```
}

/*-------------------------------------------------------------*/

/* SENSES looks at the LDR and bump sensors */

int senses(level)
{
int total;
total = level;
if ((right_eye)||(!right_bump)) total += 1;
if ((left_eye)||(!left_bump)) total -= 1;
return total;
}

int bumps(rev)
{
int hit;
hit = rev;
if ((!left_bump)||(!right_bump)) hit +=1;
else
        {
        if (hit>0) hit -=1;
        }
return hit;
}
```

## B.2 Code for simulation interface

```
/* functions for simulation */
extern int l_amp, r_amp, l_phase, r_phase;
extern int store[2500][6];
extern int index;
extern TCB tlisten, tresponse, tright, tleft;

int current_move; /* variables for interaction of move and position*/
int next_move;
long lastchange;
static int axdist[] = {250,250,250,250,251,251,251,252,252,253,254,255,256,257,258

int orientation; /* position variables */
int distance;
int alpha;
int pos;

#define length 6 /*parameters for movement calculations*/
#define swivel 7
#define inertia 500 /*stops calculation of brief changes in moves */

/*-----------------------------------------------------------------*/

/* INIT initialises the variables used in the following functions */

init()
{
/*starting point*/
orientation = 15;
distance = 250;
alpha = 15;

current_move = forward;
next_move = forward;
lastchange = gettime();
}

/*-----------------------------------------------------------------*/

/* SOUND simulates the movement of the ear circuit using a simplistic
acoustic model: it generates the syllable structure and once per
syllable updates the position estimate and thus the estimates of
amplitude and phase at the ears. */

sound()
{

if (gettime()&1024)   /*i.e. time is between syllables */
{
```

```
pos = 1;
l_amp = r_amp = l_phase = r_phase = 0;
}
else /*current values of phase and amp will be used*/
{
if (pos==1)
{
pos = 0;
position(gettime()-lastchange);
}
}
if (index<2500)
        {
        store[index][0] = 0;
        store[index][1] = l_phase;
        store[index][2] = l_amp;
        store[index][3] = r_phase;
        store[index][4] = r_amp;
        store[index][5] = (int)(gettime()>>2);
        index += 1;
        }
}

/* PHASE estimates a phase difference of 0 - 15 based on the orientation,
using (1 + cos(orient))*15/2 */
phase()
{
l_phase = ((1024 + cos(orientation))*15)>>11;
r_phase = ((1024 + cos(orientation+180))*15)>>11;
}

/* AMPLITUDE estimates different ear distances and hence amplitudes:
the ear distance is the distance + 15cm*cos(orient);
the amplitude value ranges from 15 at 25cm to 3 beyond 500cm,
decreasing by the inverse square of the distance */
amplitude()
{
int ear_distance;
/*left*/
ear_distance = distance - ((15*cos(orientation))>>10);
if (ear_distance<25) l_amp = 15;
else
    l_amp = 75/root(ear_distance);
/*right*/
ear_distance = distance + ((15*cos(orientation))>>10);
if (ear_distance<25) r_amp = 15;
else
    r_amp = 75/root(ear_distance);
}

/* POSITION estimates the current position of the vehicle depending on
movement. As these estimates take calculation
```

```
time, the execution of the rest of the program is suspended and
time is halted while it takes place (in the real robot movement and
phase and amplitude measurement genuinely occur in parallel).*/

position(duration)
{
long int halt_time;
if (duration>inertia) /* need to calculate move */
        current_move = next_move;
/*hold*/
halt_time = gettime();
suspend(&tresponse);
suspend(&tleft);
suspend(&tright);
/*position*/
if (current_move==forward) do_move(duration>>length);
else if (current_move==turnleft) do_turn(duration>>swivel);
else if (current_move==turnright) do_turn(-(duration>>swivel));
else if (current_move==back) do_move(-(duration>>length));
/*sound*/
amplitude();
phase();
/*record*/
if (index<2500)
{
store[index][0] = current_move;
store[index][1] = duration;
store[index][2] = distance;
         store[index][3] = orientation;
         store[index][4] = alpha;
store[index][5] = (int)(halt_time>>2);
}
index += 1;
/*show*/
        printf("\n\r %i %i %i",distance,alpha,orientation);
/*lay-on*/
        settime(halt_time);
resume(&tresponse);
resume(&tleft);
resume(&tright);

}


/* DO_TURN updates orientation by degrees, and adjusts it to within 0-359 */
do_turn(degrees)
{
orientation += degrees;
orientation = adjust(orientation);
}

/* DO_MOVE calculates the new position (distance,orientation and alpha) */
```

```
do_move(cm)
{
long int b, c, d;
c = cm*cos(orientation);
d = distance*1024 - cm*sin(orientation);
b = arctan((c*1024)/d);

alpha = adjust(alpha + b);
orientation = adjust(orientation - b);
distance = d/cos(b);
if (distance<0)
        {
        distance = 0 - distance;
        orientation = adjust(orientation+180);
        alpha = adjust(alpha+180);
        }
}


/*-------------------------------------------------------------*/

/* MOVE takes movement commands and stores them in next_move to
be used in position calculations when inertia time has elapsed */

move(manner)
{
current_move = next_move;
lastchange = gettime();
/* convert motor command to appropriate action */

next_move = manner;

/*if (current_move==straight)
{
if (manner==revleft) next_move = turn_left;
if (manner==revright) next_move = turn_right;
}
if (current_move==turn_left)
{
if (manner==forleft) next_move = straight;
if (manner==revright) next_move = backup;
}
if (current_move==turn_right)
{
if (manner==revleft) next_move = backup;
if (manner==forright) next_move = straight;
}
if (current_move==backup)
{
if (manner==forleft) next_move = turn_right;
if (manner==forright) next_move = turn_left;
} */
```

```
}


/*-------------------------------------------------------------*/

/* SENSES gives movements of bumpers/LDRs when obstacles are
encountered. Assumes front to +/- 90 degree coverage at 10cm (bump)
and 30cm (IR) range, based on enclosure of 4metre radius with speaker
at centre */

senses()
{
int total;
total = 0;
if ((alpha<=270)&(alpha>=180))
        {
if (distance>370)
                total+= 1;
        }
if ((alpha>=90)&(alpha<=180))
        {
        if (distance>370)
                total-= 1;
        }
return total;
}

int bumps(rev)
{
int hit;
hit = rev;
if ((alpha<=270)&(alpha>=900))
        {
if (distance>390)
                hit+= 1;
        }
else
        {
        if (hit>0) hit -=1;
        }
return hit;
}
```

## B.3 Control code

```
#include "phono.h"
#include "trig.h"
#include "mrobot.h"

/* Program being incremented towards mechanism for phonotaxis. It can
include the file "simul.h" which provides interaction with a
simulation, or "robot.h" which interacts with a Lego robot. Thus
identical mechanisms can be tested in the two systems.*/

/*------- Variables and functions used by tasks ---------------------- */

#define syll 3200 /* ISI of about 600ms */
#define ttime 1000
long int lstime, rstime;
int l_amp, r_amp, l_phase, r_phase;
int lresp, rresp;
int lan1, ran1;
int tendency;
byte signal;
int lme,lmb,rme,rmb,backup;

/*------- Array used to store data ---------------------------------- */
int index, item, run;
int store[2500][6];

/* explicit initialization is needed because variables are not reset when
tasks are killed and reforked by onoff switch */

initialize()
{
    index = 0;
    item = 0;
    lan1 = 0;
    ran1 = 0;
    tendency = 0;
    signal = 0;
    lme = lmb = rme = rmb = backup = 0;
    init();  /*initializes simulation or robot variables */
    lstime = rstime = 0;
}

/* RESPONSE estimates the response amplitude of the combined waves
using combined^2 = ampA^2 + ampB^2 - 2*ampA*ampB*cos(phase-difference) */
response()
{
int sum, factor;
long temp;
sum = l_amp*l_amp;
```

```
sum += r_amp*r_amp;  /* because l_amp*l_amp + r_amp*r_amp doesn't work */
factor = l_amp*r_amp;
temp = factor*cos(12*l_phase);
temp = temp>>9;
lresp = root(sum - (int)temp);
temp = factor*cos(12*r_phase);
temp = temp>>9;
rresp = root(sum - (int)temp);
}

/*========== TASKS ==================================================*/

/* LISTEN updates the reading from the ears every 35ms,
calculates the response value, then applies the low-pass filter
"y(t+1) = 7*y(t)/8 + x" to emulate the neural response */

TCB tlisten;

listen()
{
long last, temp;


    last = gettime();
    for (;;)
    {
        temp = gettime();
     while ((gettime()-last)< 150); /* do following every 35ms*/
last = gettime();
sound(); /*read phase and amplitude values*/
response(); /*calculate response values*/
lan1 = ((7*lan1)>>3) + lresp; /*neural filter*/
        ran1 = ((7*ran1)>>3) + rresp;
   if (lan1>16) lan1 = 16; /*threshold*/
if (ran1>16) ran1 = 16;

      }
}

/*====================================================================*/
/* LEFT and RIGHT coordinate motor signals generated by response from AN1,
AVOID handles response to sensed obstacles. Note that all movements are
called only when they change, which hopefully avoids motor conflicts.
Turns in response to sound are followed by pauses */

TCB tleft;
TCB tright;

left()
{
long int now;
move(forward);
```

```
for(;;)
{
/* left motor forward */
        while (lstime<gettime())
        {
        resched();
        }
leftstop; /* turn for length ttime */
now = gettime();
while (gettime()<(now+ttime))
{
resched();
}
move(not); /* pause till effect of lstime elapses */
now = lstime;
while (gettime()<lstime)
{
if (lstime>now) /* turn if lstime increases */
{
suspend(&tright);
now = lstime;
                        leftstop;
        rightfor;
}
resched();
}
resume(&tright);
leftfor;
}
}

right()
{
long int now;
for (;;)
{
        while (rstime<gettime())
        {
        resched();
         }
rightstop;
now = gettime();
while (gettime()<(now+ttime))
{
resched();
}
move(not);
now = rstime;
        while (gettime()<rstime)
        {
if (rstime>now)
{
```

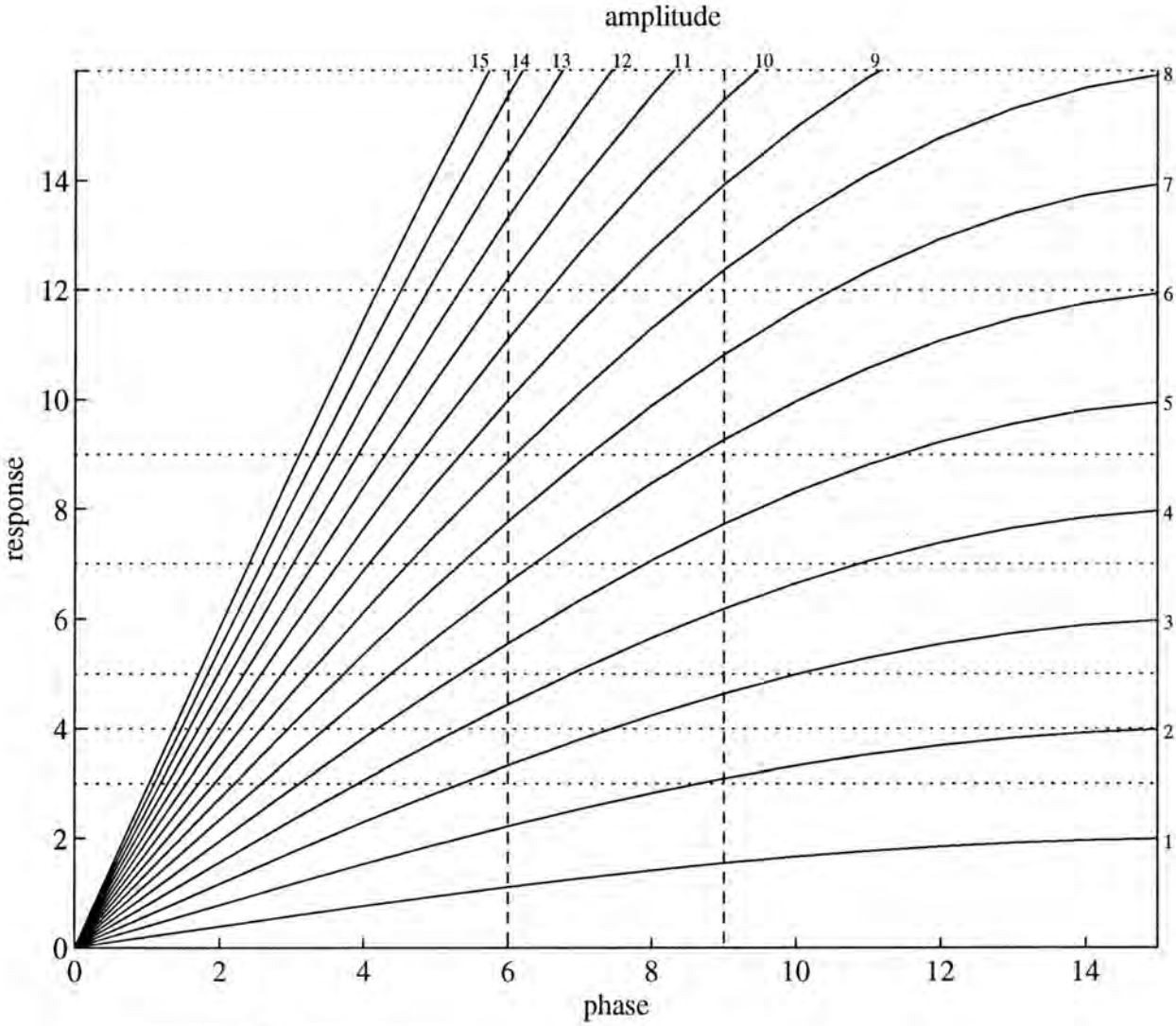
```
suspend(&tleft);
now = rstime;
                        rightstop;
leftfor;
}
        resched();
        }
resume(&tleft);
rightfor;
}
}

TCB tavoid;
avoid()
{
for (;;)
        {
        if (backup>0)
                {
                move(back);
                while (backup>0)
                        {
                        backup -=1;
                        resched();
                        }
                move(forward);
                }
        resched();
        }
}

/*=========================================================================*/

/* RESPONSE determines the movement commands which are based on the
level of an1 */

TCB tresponse;

response()
{
int lefton, righton;    /* used to distinguish onsets */
int level, sample, rev ; /* used to sample sensors */
sample = rev = 0;
level = 20;
lefton = righton = 0;
for (;;)
    {
    if (lefton|righton)
        {
        if (lan1<8) lefton = 0;
        if (ran1<8) righton = 0;
        }
```

```
    else
        {
        if ((lan1>15)&&(ran1<16))
                {
                lstime = (lstime>gettime())?(lstime+syll):(gettime()+syll);
                lefton = 1;
                }
        else if ((ran1>15)&&(lan1<16))
                {
                rstime = (rstime>gettime())?(rstime+syll):(gettime()+syll);
                righton = 1;
                }
        }
/* check sensors */
    rev = bumps(rev);
    if (rev>1)
        backup += 10;
    if (sample<10) /* continue sampling IR */
        {
        sample = sample + 1;
        level = senses(level);
        }
    else /* if level has moved from 20 indicate turn */
        {
        if (level>20)
             {
             lstime = gettime()+1000;
        }
        else if (level<20)
        {
             rstime = gettime()+1000;
        }
level = 20;
sample = 0;
}
    }
}

/*============================================================*/
main()
{
long d;

lightson;
for (;;)
{
printf("\nReady to start\r");
hshk(0);
while onoff; /*wait for start switch*/
initialize();
correlate(); /*lights flash before moving*/
        fork(&tlisten,listen,USER,256);
```

```
        fork(&tresponse,response,USER,256);
        fork(&tleft,left,USER,256);
        fork(&tright,right,USER,256);
        fork(&tavoid,avoid,USER,256);
        for (d=0;d<20000;d++);  /*switch clear*/
while onoff
                {
if (index>2000)
goto Stop;
if (!pause)
{

suspend(&tlisten);
suspend(&tresponse);
suspend(&tleft);
suspend(&tright);
suspend(&tavoid);
turnoff();

for (d=0;d<50000;d++);  /*switch clear*/
printf("\npaused\r");
while pause  /*wait for pause switch*/
if (!onoff) goto Stop; /*watch for stop switch*/

resume(&tlisten);
resume(&tresponse);
resume(&tleft);
resume(&tright);
resume(&tavoid);
turnon();


printf("\nrestarted\r");
for (d=0;d<50000;d++);  /*switch clear*/
}
else
                        resched();
}
Stop: printf("\nended\r");
kill(&tlisten);
kill(&tresponse);
kill(&tleft);
kill(&tright);
kill(&tavoid);
lightsout;
while pause; /* wait for pause switch then print data */
        hshk(1);
        run = index;
printf("\n\r%i",run);
        for (index=0;index<run;index++)
{
printf("\n");
```

```
for (item=0;item<6;item++)
                        printf("%i ",store[index][item]);
}
lightson;
    }
}
```

# Appendix C




## Latency Differences:

Each solid line on the graph represents a different *amplitude* value (from 0 – 15), and shows the *response* level for different values of *phase* (0 – 15) at that amplitude. Different *response* levels result in a different number of steps (latency) for *an* to reach threshold (see Figure 5-9): the horizontal dotted lines show the *response* levels (3,4,5,7,9,12 and 16) at which there is a change in number of steps (the latency) to threshold (9,7,5,4,3,2 and 1, respectively). Thus if two given *response* levels lie on either side of one of these lines, there will be a latency difference between them.

Because of the symmetry of phase calculation, if there was no noise in the signal, the *right phase* would always be $15 - left\ phase$; even with noise this is approximately true. Thus a difference in phase of more than 2 would mean *right phase* = 6 and *left phase* = 9, or vice versa (the vertical dashed lines on the graph). It can be seen that for amplitudes from 2 – 15 the response values at 6 and 9 fall on opposite sides of a latency line.